IMPROVING THE KINEMATIC CONTROL OF ROBOTS WITH COMPUTER VISION

by

J. Barry Fallon

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MECHANICAL ENGINEERING

APPROVED:

Charles F. Reinholtz, Chair
Mechanical Engineering

A. Lynn Abbott
Electrical Engineering

Sanjay G. Dhande
Mechanical Engineering (IIT, Kanpur)

Michael P. Deisenroth
Industrial and Systems Engineering

Layne T. Watson
Computer Science

July, 1995

Blacksburg, Virginia

Key Words: Robotics, Computer Vision, Tracking, Pose, Control

IMPROVING THE KINEMATIC CONTROL OF ROBOTS WITH COMPUTER VISION

by

J. Barry Fallon

Charles F. Reinholtz, Chair

Mechanical Engineering

(ABSTRACT)

This dissertation describes the development and application of a computer vision system for improving the performance of robots. The vision-based approach determines position and orientation (pose) parameters more directly than conventional approaches that are based on kinematics and joint feedback. Traditional robot control systems rely on kinematic models, measured joint variables, knowledge of objects in the workspace, and the calibrated robot base pose to correctly position and orient a tool. Since this conventional approach involves a large number of parameters, unacceptable pose errors may accumulate. In contrast, the vision system approach uses images from a tool-mounted camera and geometric knowledge of objects in the workspace to accurately track and determine the end-effector pose. This approach is advantageous because the camera directly observes the parameters of interest (position and orientation of the robot tool with respect to the work-piece) during the positioning process.

The vision approach is verified and its utility demonstrated by increasing the automation and accuracy of computer controlled robots used in the nuclear service industry. The overall solution strategy involves tracking and pose determination. Tracking is used as a coarse positioner and to verify the toolhead position prior to performing crucial servicing operations. Pose determination is used to calibrate the base location of the robot, verify the tool pose for insertions, and compute a precise correction if necessary.

The major contributions of this work lie within its comprehensive treatment, which begins with theoretical modeling and follows through to the details of application. Specific contributions are made in the areas of robotics, image processing, calibration, tracking, pose determination, kinematic control strategies, and nuclear service operations. Performance results from laboratory experiments and actual field testing have been encouraging. The vision-based strategy offers robustness to the conventional error stack-ups encountered in robotics and promises to improve the accuracy, flexibility, and cost of both specialized and general-purpose robotic systems.

Dedicated to my parents

Patricia H. and Robert E. Fallon

for committing over 45 years of their lives
to raise and educate their ten children

Acknowledgments

After reflecting upon this dissertation and my four years of graduate study at Virginia Tech, I feel indebted to the many people who have sincerely helped me accomplish this work. The contributions each has made to both my personal and professional development are greatly appreciated:

-Charles Reinholtz, my graduate advisor and Jeri Reinholtz

-Other members of my graduate and examining committees: Lynn Abbott, Sanjay Dhande, Layne Watson, Mike Deisenroth, Bob Salerno, and Steve Shooter

-B&W Nuclear Technologies for supporting this work with three research grants, particularly Bill Glass and Frank Klahn

-Virginia's Center for Innovative Technology for supporting this work with three matching grants, particularly Wayne Hawkins

-National Science Foundation for also supporting this work

-Virginia Tech's Multimedia Lab, particularly Gordon Miller

-Duke Power's McGuire and Catawba nuclear power stations

-Robotics and Mechanisms Group at Virginia Tech, particularly: Arun, Jimmy Asher, Joe Calkins, Anne Carrithers, John Collier, Robert Gage, Anil Garg, Amber Jenkins, Hong-you Lee, Allen Maples, Curtis McKim, Vittorio Provera, John Stulce, Paul Tidwell, Ryan Vallance, Ravi Voruganti, and Jeff Wolfe.

-Mechanical Engineering Department, particularly Lynne Ellis, the mechanical and electrical shops, and Doug Nelson.

-Industrial and Systems Engineering Department for the loan of equipment

-Dan Farkas

-David Coe

-My family: Robert and Patricia, Cindy, Shauna, Joyce, Bob, Patty, Billy, Brian, Brandon, and Maura

- Lee and Susan Thompson, and Anne Gosling

Table of Contents

List of Figures

List of Tables

Chapter 1

1. Introduction

This dissertation describes the development of a computer vision system for improving the kinematic control of robots. The vision approach is advantageous because the camera directly observes the parameters of interest, namely, the position and orientation of the robot tool with respect to the workpiece. The vision system is successfully applied to increase the automation and accuracy of nuclear service robots. This introductory chapter provides background and motivation for this work, a summary of the approach and some important design considerations, the author's contributions, and finally, a preview of the remaining dissertation.

1.1 Background

Process control is important in manufacturing and service industries as well as many other aspects of modern society. A process is a systematic series of actions directed at a desired end. At its most basic level, process control involves four steps:

1. establishing the goals of a process,
2. determining the parameters that affect the process,

3. measuring the error between the current process state and the goals during the process, and

4. affecting the process in a positive way so that the goals are more precisely obtained.

This strategy is a powerful tool, particularly when high performance is sought from a system. These ideas can be widely applied in the field of robotics in areas such as welding, spray painting, and general positioning tasks.

In the process of robotic spray painting, the desired goal is to obtain an attractive, protective, and durable coating. Often the process is monitored by sensing parameters such as air pressure, paint flow, ambient temperature, paint viscosity, and humidity. Additionally, the position, orientation, and geometry of the workpiece must be known and the motions of the robot preprogrammed. Following the painting process, the quality of the finish is inspected with manual, optical, or computer vision methods. The problem with this approach is that deviations from the actual process goals are not monitored directly during the process, rather they are inspected after the process is completed. During the process, other parameters that tend to yield acceptable results are monitored because they are more easily measured with existing technology. If problems develop, the quality of the product diminishes until the inspection system collects consistent results and the process equipment is adjusted or repaired. Because of the time lag between the painting process and inspection, many products may be produced with unacceptable quality before the problem is finally corrected. If deviations could be monitored directly during the process and the process continuously adjusted, then higher product quality could be obtained. Additionally, post-process inspection and product rejection would become unnecessary, and the efficiency of the process as a whole would increase.

In the process of robotic arc welding the desired goal is to obtain a uniform, correctly sized weld bead with proper fusion and penetration. Many robotic systems depend on

preprogrammed knowledge of the workpiece position, orientation, and geometry. More sophisticated systems monitor the arc current or use vision systems to track the weld joint. Once again, however, deviations from the process goals are not measured directly and some level of inspection must be performed after the process is complete to ensure quality.

A third example and the primary focus of this dissertation involves a large class of tasks where the robot end-effector must be positioned and oriented precisely relative to a workpiece. Conventional robot control strategies rely on kinematic and dynamic theory to successfully position and orient a robot end-effector. Kinematic theory uses geometric models of the robot, work-cell, and workpiece to plan accurate and efficient paths. In addition to the desired pose (position and orientation) of the end-effector, path planning also considers singularities, obstacles in the workspace, and mechanical limits of joints and cables. Dynamic theory can be used to develop models that minimize overshoot, settling time, and steady-state errors during joint actuation. Under this approach, substantial pose errors can result from the accumulation of modeling errors including the following:

1. position and orientation of the robot base,
2. link and joint geometry (length, offset, and twist),
3. initialized and relative joint variables such as length and angle,
4. end-effector tool dimensions,
5. workpiece position and orientation,
6. knowledge of the work-cell environment, and
7. deformations in robot components such as the support base, links, joints, mechanical connections, and end-effector tools caused by bending, shear, axial, and torsional loads.

The items listed above reflect robot construction, installation, and operation. Even if the robot is manufactured with high tolerances and subjected to small loads, substantial pose errors can still accumulate because of variations in the workpiece or work-cell environment. Furthermore, error accumulation is typically greater for large manipulators, mobile robots, and systems that are remotely installed in hazardous environments such as the long-reach, parallel manipulator proposed by Salerno (1992 and 1993) and illustrated in Figure 1.1.

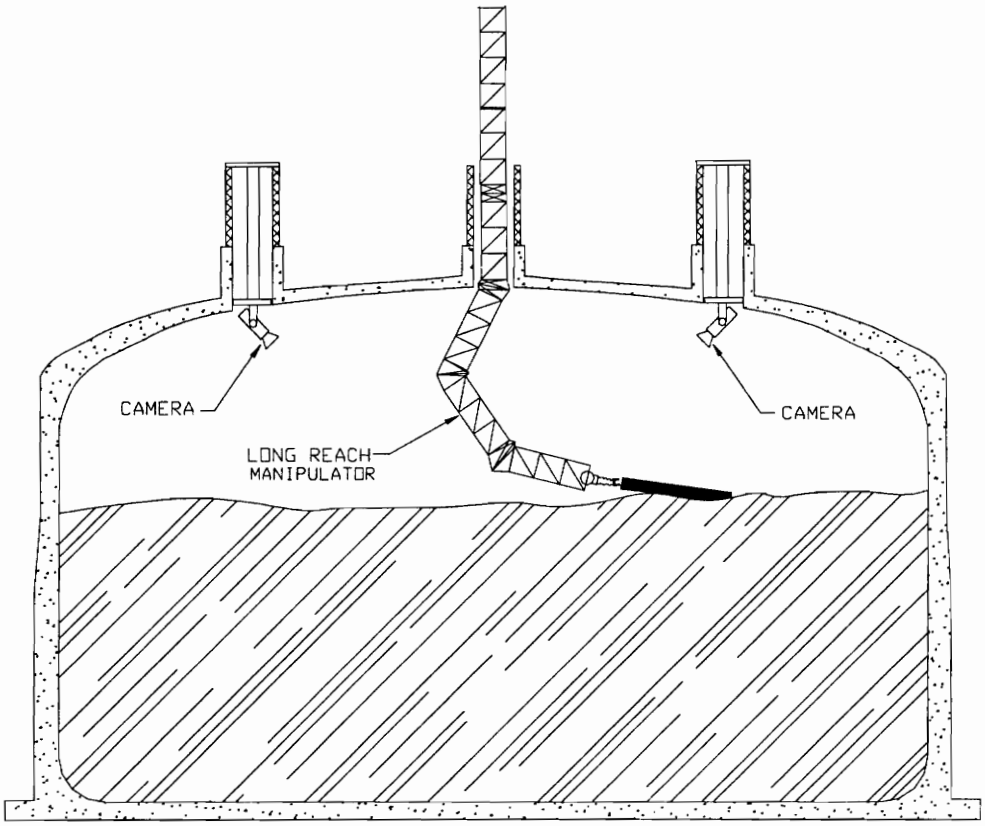


Figure 1.1: Long-Reach Manipulator Remediating a Nuclear Waste Storage Tank

The subject of this dissertation is improving robot control by directly sensing the relative pose (position and orientation) between an end-effector and workpiece. A variety of technologies can be employed to improve pose sensing, including laser interferometry, acoustic time-of-flight, and computer vision. Such methods offer robustness to the conventional error stack-ups encountered in robotics.

1.1.1 Direct Pose Sensing Approaches

Laser interferometry offers accurate, precision measurement in working volumes as large as 28.3 m^3 (1000 $ft.^3$). Brown et al. (1986 and 1988) give a detailed description of one such system which consists of a laser tracker, one or more retro-reflectors on a moving object, a laser position sensor coupled to servo motors on the tracker, and an interferometer. A six axis pose measuring system employs multiple trackers to measure the position of three retro-reflectors. Each tracker is encoded to accurately measure the laser's azimuth and elevation pointing angles, while interferometry is used to measure the distance along the beam's length. Positions are reportedly measured to within 8 microns (0.0003 in.) in 3.05 m (10 ft.). Another laser measurement system consists of a mobile detector and two or more scanning heads (Spatial Positioning Systems, 1995). Each head produces two moving planes of light that scan a large workspace. The detector's position defines the intersection of the light planes, which then gives the detectors position in the workspace. Although lasers-based systems offer high performance, they generally involve many complex components and are expensive.

Acoustic systems use the time of flight of acoustic signals to determine the distance between acoustic source/receiver pairs. Expensive tracking hardware is not required for most acoustic systems because sound can propagate in all directions. Tidwell et al. (1993) describe such a system for measuring the pose of an underwater mobile robot in a nuclear reactor vessel. Accuracies of 20 mm (0.8 in.) are possible when operating at ranges up to 100 m (328 ft.) in homogeneous underwater environments. Another acoustic system, used in air, consists of a hand-held pointing device and a fixed digitizing frame that contains three or more sound receivers (Science Accessories Corporation, 1995). Two sound emitters, precisely mounted on the pointing device, enable the tip of the pointer to be located within the digitizer's work-space. Accuracies of 0.5 mm (0.020 in.) are reported in volumes of 2.4 cubic meters (85 $ft.^3$). Although acoustic systems are less expensive than those based on laser interferometry, they are not nearly as accurate and are sensitive to acoustic interference as well as variations in the propagation medium.

Computer vision offers many approaches to sense the pose of objects, including stereo, multi-view monocular, and single-view monocular imaging. Stereo vision utilizes corresponding points in images obtained from two cameras with a known relative pose. The advantage of this method is that specific prior knowledge about the object geometry is not required. The disadvantages are that two cameras must be used and establishing a correspondence in the two images can be difficult. Multi-view monocular imaging is similar to stereo, but uses a single camera that undergoes known relative motion to create multiple images of the same object. Single-view monocular imaging requires only a single view from one camera and can be used if prior knowledge of the object geometry is available. In addition to geometric information that is inherently available in many manufacturing and service settings, targets and lasers can further encode the scene. In some applications, computer vision can provide pose accuracies that are substantially better than acoustic systems, and at significantly lower costs than laser systems.

In the work described here, single-view monocular imaging is further divided into a two-step approach. The first step uses a fast and computationally efficient method that locates and tracks features as the end-effector moves in the workspace (Fallon et al., 1994b and 1995). Once the workpiece is located and a coarse relative position between it and the end-effector obtained, a more precise static pose is computed (Fallon et al., 1994a).

1.1.2 Motivation

The motivation for this work evolved from recent robotic development projects in the nuclear service industry. Currently, the United States is the world's largest producer of nuclear-based electric power with 109 reactors producing nearly 100,000 megawatts. This amounts to about 20 percent of the total domestic electrical power consumption and enough electricity to illuminate 1 billion 100 watt light bulbs. France is the second largest with 56 reactors producing about 57,000 megawatts. The United States operates over 25% of the world's commercial reactors along with 158 additional reactors to power

U. S. Navy vessels. About two-thirds of the commercial units are pressurized water reactors (PWRs), while the remaining one-third are boiling water reactors (BWRs). Each PWR has between 2 and 4 steam generators which are inspected and repaired every 15-18 months while the reactor is being refueled.

Steam generators, such as the one shown in Figure 1.2, are large shell-and-tube heat exchangers consisting of several thousand tubes. Tube diameters range from one to two centimeters. Used in conjunction with pressurized water reactors, steam generators produce steam which powers turbines and generates electric power. Water is continuously heated in the nuclear reactor and circulated through the primary inlet, tubes, and primary outlet of the steam generator before returning to the reactor. As the primary water flows through the tubes, heat is transferred across the tube wall to the secondary water which is converted to steam. The steam generator tubes form a boundary between the high pressure, primary water from the reactor and the secondary water and steam, which is under relatively lower pressure. To maintain the pressure boundary, electric power utilities spend about three of the total six to eight week refueling outage servicing steam generators. It is important to minimize the inspection and repair time and also the amount of radiation exposure encountered by workers. The cost of outage time is estimated as the cost to replace the electricity lost by the reactor shut-down. In many cases, such costs approach two million dollars each day a reactor is out of service.

Steam generators are serviced by installing a portable robot though a 0.4 m (16 in.) diameter manway into the inlet and outlet plenums below the tube bundle (Tidwell et al., 1991). The robot is then used to position a variety of servicing toolheads at the open ends of the tubes. All of the toolheads contain a small video camera that gives the robot operator a view of the tool and a few dozen tube-ends. Potential problems in the tube walls are diagnosed by passing eddy-current probes through the U-shaped tubes, which can be as long as 25 meters. Deficient tubes are either repaired by sleeving or removed from service by plugging both ends.

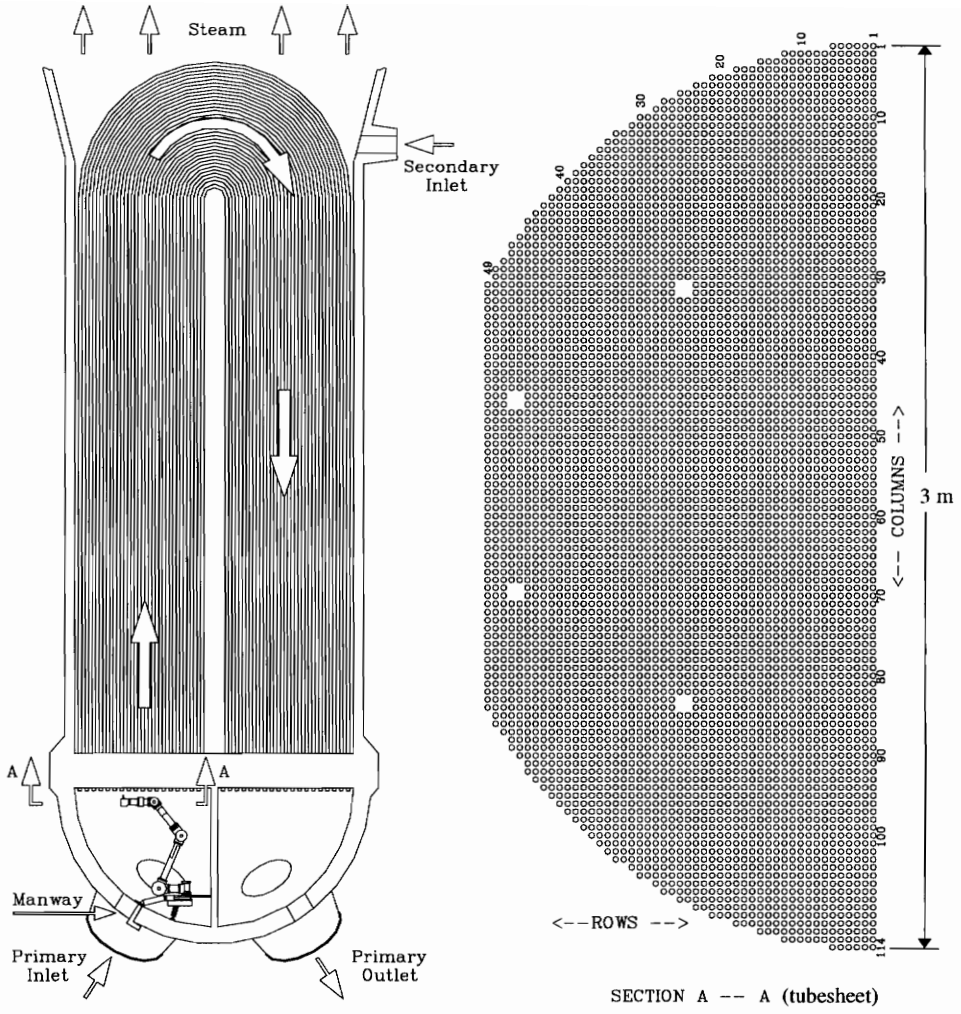


Figure 1.2: Steam Generator

Steam generator robots are classified as limited access workspace (LAWs) robots because they must service large workspaces (up to 11,000 tubes) while being installed through relatively small portals (Shooter et al., 1992). In addition to LAWs constraints, these robots must also be lightweight for portability and still be capable of lifting heavy tool loads. The Cobra steam generator robot, pictured in Figure 1.3, weighs only 50 kg (110 lb.), but can lift over 68 kg (150 lb.) at a full extension of 1.94 m (77 in.) (B&W Nuclear Technologies, 1995). By comparison, the popular Unimation PUMA 560 robot weighs 63 kg (140 lb.) and is rated for a maximum load of 4 kg (8.8 lb.) at its full reach of 0.92 m

(36.2 in.). The accuracy requirements for Cobra are quite high considering its strength-to-weight ratio and remote mounting. An absolute positional accuracy of ± 0.5 mm (0.020 in.) in the horizontal tubesheet plane is desired to calibrate the robot base position and ± 1.25 mm (0.050 in.) is desired for tool positioning.

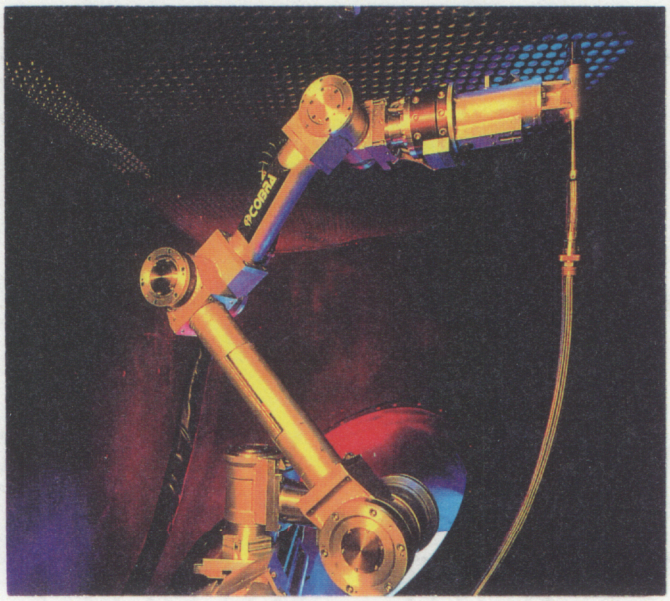


Figure 1.3: Cobra Steam Generator Robot

In the current control method for robots such as Cobra, kinematic equations are solved to determine the angles each of the joints must move through to position a tool at a desired tube location. These angles become reference inputs and, together with joint synchro sensors, establish an error signal to the PID (proportional-integral-derivative) controller of each joint. When the joints move through the appropriate angles, the error signals to the controller and correspondingly, robot motion ceases. The position of the tool with respect to a particular tube of interest is estimated by knowing the steam generator and robot geometry and measuring the joint angles. Referring to Figure 1.4, this is equivalent to tracing a path from the tube of interest, across the tubesheet, around the bowl, up the mezzanine support base, through each joint and link of the robot, and finally through the tool to the tool-tip.

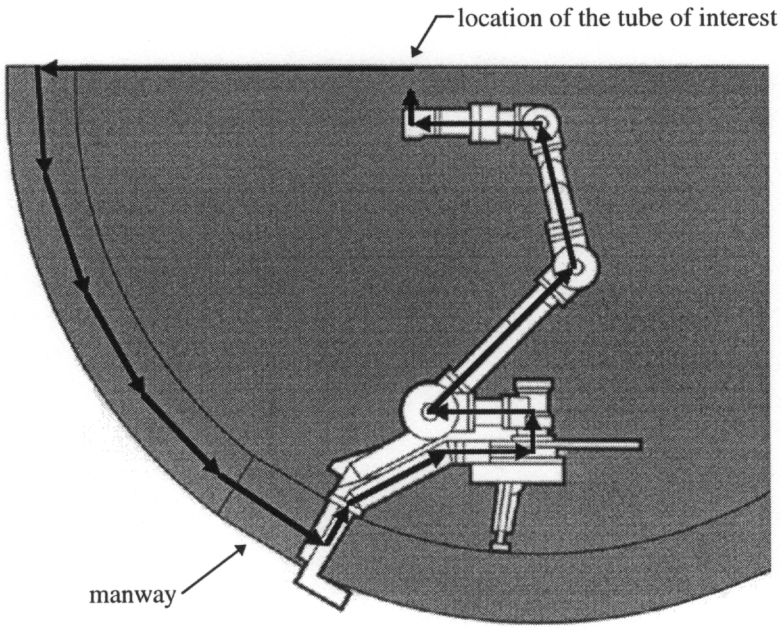


Figure 1.4: Conventional Error Stack-up

If the robot worked in a perfect manner and operated in a perfect environment, then the tool would reach the desired tube center. However, because small errors propagate from many sources as discussed previously, a tool position error results. Since the error signal to the controller has ceased, a manual correction may have to be applied before a tool can be successfully inserted into a tube. This is an example of indirect process control which uses an indirect reference input and feedback (kinematically solved and measured joint angles) rather than a direct input and feedback such as the position of the toolhead with respect to a desired tube center in the bowl coordinate system. Thus, the control system may not be able to automatically position the toolhead at a desired tube location with sufficient accuracy for insertion in all cases.

Currently, error accumulation is minimized by precisely measuring intrinsic robot and tool parameters off-line, calibrating robot base pose parameters with a calibration process after the robot is installed inside the steam generator (Shooter and Reinholtz, 1992), and on-line compensation for deflections in the robot and base mounting (Calkins, 1994). The robot base pose parameters consist of six numbers: three describing the position of the base in the tubesheet coordinate system and three for orientation. Two orientations

are known beforehand since the robot base is initially mounted level with the tubesheet. To calibrate the remaining unknown parameters, an operator tele-operates the robot to three or more regions in the generator and docks the tool-tip on the ends of specific tubes. With the known robot geometry coupled to the known tubesheet geometry, the base parameters can be computed and then used by the computer for automatic tool positioning throughout the servicing outage. During the calibration, however, manually docking tubes in the tele-operated mode requires a high degree of operator skill and is time-consuming.

Even with the careful approach described above, a second, independent measure of tool position is needed to ensure quality. The second method consists of a manual count of tubes from known tubesheet reference points. By viewing the tubesheet through the camera mounted on the robot tool, the operator must verify the exact row and column number address of the tube being serviced. Since the video image only shows about two dozen of the thousands of tubes in a tubesheet, it is generally impossible to determine the address of any tube in a randomly selected image. The operator must start viewing with the robot at some landmark tube location, such as a plugged tube or one near the edge of the tubesheet. From this landmark tube, row and column numbers are counted as the robot moves under the tubes to the tube of interest. Typically, verifications are performed after inspecting sets of 20 tubes and before any tube is repaired. Repair activities occur only after tubes are re-inspected to confirm flaws. The tubes to be repaired are often marked and yet another verification performed on the marked locations. The guiding principle is to check and double check to avoid costly mistakes.

Using the existing tool-mounted camera, computer vision can automate the robot base calibration and tool position verifications, as well as offer an alternative approach to the overall control method. Vision provides a more direct measurement of the parameters of interest, namely, the position and orientation of the tool with respect to a particular tube than the existing control approach. The overall solution strategy involves tracking and pose determination. Tracking is used as a coarse positioner and to verify the toolhead position

prior to performing crucial servicing operations. Pose determination is used to calibrate the base location of the robot without docking tubes, verify the tool pose for insertions, and compute a precise correction if necessary.

1.1.3 Overview of the Vision Approach

Figure 1.5 illustrates the direct sensing approach using a single-view, monocular image from a tool-mounted camera. The goal is to recover the tool-to-tubesheet pose relationship, as shown in the figure. First, vision methods are used to characterize camera, optics, and digitizer parameters before the camera is installed on the robot tool. After the camera is securely attached to the tool, vision is again used to calibrate the tool-to-camera pose relationship. The camera-to-tubesheet pose is determined in two steps, both of which rely on precise knowledge of the tubesheet geometry. The first step is a fast and computationally efficient method that locates and tracks features as the tool moves in the workspace. Once the tube of interest is located and a coarse relative position between the end-effector and the workpiece obtained, the second step computes the precise pose. The camera-to-tubesheet pose is then used with the calibrated tool-to-camera pose to obtain the tool-to-tubesheet pose, which is the desired result.

Tracking the toolhead is accomplished by continuously tracking two tube-ends in the toolhead camera video image. At least two tube-ends must be tracked because the axes-symmetric nature of a single tube-end causes ambiguity of the viewing direction. As the robot moves, tubes continually enter, move around, and leave the image, so the system must continually look for new tubes to track. The tracking process is divided into three main tasks:

1. computing the centers of two tube-ends being tracked in the current video frame, given their locations in the previous frame as estimates (the camera acquires 30 frames/second);

2. replacing the tube-ends currently being tracked with better tube-ends as they become available; and

3. using the tubes currently being tracked to estimate the location of some desired tube in the image (e.g., a target tube).

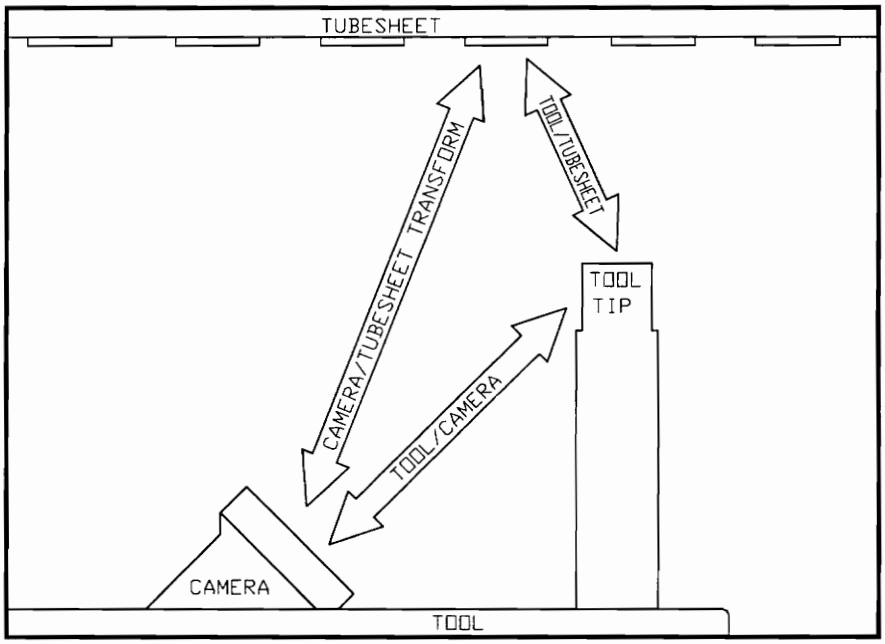


Figure 1.5: Single-View, Monocular Imaging Approach

The precise pose is then determined using the following process:

1. establishing a correspondence between geometric features known in object space and their image (i.e., matching images to known objects);

2. using the known tubesheet geometry in object space together with measurements on the image plane to determine the spatial transform between the camera and tubesheet coordinate systems; and

3. computing the spatial transform between the tool-tip and the camera by initially applying the above procedure with the tool docked on a tube-end.

1.1.4 System Design Considerations

Developing computer vision technology for use in robotics entails many practical and theoretical difficulties. Vision systems must function with minimal human interaction, high reliability, and robustness in demanding environments over which there may be little control. In the present application, lighting and tool camera viewing angle constraints together with the oxidation properties of steam generator tubesheets produce low contrast images. These images must then be acquired with low-cost CCD (charged-coupled device) cameras and wide angle lenses having optical focal lengths as low as 5 mm. In addition, the video signals are multiplexed and transported over 300 meters from the steam generator containment building to the service control trailers. The video signals may contain substantial noise due to electromagnetic interference, poor connections, and radiation, which degrades CCD camera sensors. Locating the tube-end of interest in an image is further complicated, because it may be partially occluded by the end of the tool-tip.

Despite these challenges, the system must be cost effective and consistent with current trends to reduce personnel radiation exposure. The system set-up, operation, and maintenance procedures must be easy to accomplish and must require minimal hardware due to space limitations in the robot control trailers and particularly inside the steam generator. Finally, the system must be compatible with many different robots, toolheads, and steam generators found in both domestic and foreign power plants.

1.2 Contributions

The major contributions of this work lie within its comprehensive treatment, which begins with theoretical modeling and follows through to the details of application. The present work builds upon previous efforts by others in various fields and makes specific contributions which include the following:

1. Formalizing the task with a process control approach, enumerating robotic error stack-ups, and examining various direct pose sensing technologies.

2. Studying lighting in the scene and developing specialized toolhead lighting to obtain high image contrast which simplifies image processing, improves reliability of tracking, and increases the accuracy of pose determination.

3. Developing a three-tier camera calibration system. The first tier uses specially developed hardware and a well-known comprehensive method to solve for internal system-specific parameters of the camera, lens, and image acquisition board. The second tier utilizes geometric knowledge of the working environment to solve for the camera external pose parameters. The third reduces the calibration problem based on a sensitivity analysis, so specialized calibration hardware is not required.

4. Developing a tracking system based on methodology previously proposed in the literature. The present work develops numerous improvements to refine and expand tracking which include independent processing of interlaced image data fields; specialized edge operator; comprehensive 3-D mathematical imaging, tracking, and pose estimation models; integration with the robot control system; strategy for user interaction; and scalability depending on the system requirements and available computational power.

5. Developing a pose determination system and demonstrating its use in non-contacting robot base calibration and guiding the robot tool in critical insertion tasks. Distinguishing characteristics include: reconstructing projected centers of circular image features which maximizes accuracy and resists blooming characteristics of CCD sensors; using a well-known closed-form method to seed a much more powerful approach in which the spatial transform

parameters are optimized directly; and integrating pose determination with calibration and tracking to form a useful feedback system for robot control.

6. Presenting results from error, sensitivity, and performance assessments.

7. Briefly examining kinematic control strategies in the context of their practical benefits.

1.3 Preview

The remainder of this dissertation is organized as follows. First, relevant literature is reviewed in the areas of camera calibration, optics, computer vision, robotics, control, mathematics, and statistics. Next, the system hardware components and software architecture are described. A general model is then presented that describes the geometric mapping from features in object space to CCD camera sensors and computer frame memory. Internal and external vision system calibrations are discussed, both in general and in the context of the present work. Well-established, comprehensive algorithms are used to solve for the internal system parameters, while specialized algorithms are developed to determine the external (camera pose) parameters and a reduced internal parameter set. The tracking and pose determination systems are then described in detail. Constraints and assumptions are discussed along with results from computer simulations and sensitivity analyses. The systems are integrated, tested in controlled laboratory experiments, and field tested with several different robot systems. Finally, performance results are given along with conclusions and recommendations for future work.

Chapter 2

2. Literature Review

The literature review is divided into five areas which are relevant to the work described in this dissertation: camera calibration and optics; computer vision; robotics, control; and mathematics and statistics.

2.1 Camera Calibration and Optics

Lenz and Tsai (1988), Tsai and Lenz (1988) and Tsai (1987) offer a two-stage technique to calibrate the internal and external parameters of off-the-shelf TV cameras. Their algorithms enable standard cameras to be used for high accuracy geometric measurements. The external parameters consist of the position and orientation of the camera with respect to a calibration template. The internal parameters consist of the effective focal length, horizontal scale factor, image center, and lens distortion coefficient. Their method simplifies the large scale nonlinear problem with a radial alignment constraint. This constraint is basically an observation that the lens distortions can be divided into two components, radial and tangential. If the tangential portion is negligibly small (as for many cameras), then the distortion can be treated as radial only. Their work gives a critical review of other popular methods, establishes a theoretical

framework, describes the algorithms, and presents experimental results. Further details on optical systems and lens distortion can be found in Driscoll and Vaughan (1978), Longhurst (1967), Nussbaum and Phillips (1976), and Slyusarev (1984). In the present work, specialized hardware was developed and used with the Lenz and Tsai algorithms to determine internal vision system parameters. New algorithms were then developed to determine the camera external parameters as well as a reduced set of internal parameters which can be used for tracking and do not require the specialized hardware.

2.2 Computer Vision

Ballard and Brown (1982), Gonzalez and Woods (1992), Haralick and Shapiro (1992), and Horn (1989) provide excellent general texts in the areas of image processing and computer vision. Topics of particular interest to the present work are image formation, perspective transformations, noise filters, edge detection, connectivity, least-square curve fitting, and motion.

Wolfe et al. (1988 and 1991) provide a good discussion of imaging geometry and the solutions of the perspective three-point problem. The solution of this problem enables three noncollinear object points to be used in computing the position and orientation of the camera from the points. Wolfe shows that of the eight total solutions, four positive, real solutions are possible, but more commonly two are encountered. Graphical representations of the solution spaces are also presented.

Merritt (1949) gives an explicit solution to the three-point resection problem investigated by Wolfe. Linnainmaa et al. (1988) use pairs of triangles to remove the ambiguity of the three-point resection problem and solve for the pose of three-dimensional objects. Zeng et al. (1993) also use sets of triangles to register the location of a laser velocimeter with respect to a structure in modal analysis experiments.

Haralick (1989a, 1989b) and Haralick et al. (1987, 1988, and 1991) derive a variety of inverse perspective projection relationships that can be used to reconstruct pose from single monocular camera views. Two cases of particular interest are planar parallelogram and triangle features. As discussed by Wolfe, Haralick shows that eight solutions are obtained for the triangle (three-point resection) problem. The parallelogram, however, yields a unique solution for the 3-D location of the four corner points. Hung et al. (1985) also offer a closed-form approach for the perspective projection of four points, but for the more general quadrangle case. Photogrammetry literature is also an important source for similar information in such texts as Gosh (1988), Merritt (1958), Slama (1980), and Wolf (1974). In the present work, closed-form methods are used to seed a much more powerful approach in which the pose parameters are determined by directly optimizing the spatial transform.

Ellis et al. (1992) give a method to detect ellipses in digital images. Rothwell et al. (1992) and Mundy and Zisserman (1992) show how to recover the position and orientation of a pair of coplanar conics from a single perspective image. A number of methods are described, including a simple four-point back projection model and the use of transformation invariants. Safaee-Rad et al. (1990, 1991, 1992a, and 1992b) also investigate quadric curves under perspective projection and the 3-D location of circular features in an image. In the present work, projected centers of circular image features are reconstructed to maximize accuracy and resist the blooming phenomenon of CCD sensors.

Brem and Nandhakumar (1993 and 1994) propose an efficient tube counting method for steam generator applications. The robot operator must first initialize the vision system by establishing several parameters including the centers of three tubes on the image plane. Tubes are tracked by finding the new center locations in subsequent frames. When the tubes that are currently being tracked move toward the edge of the frame, a jump is performed to tubes that are closer to the image center. This method includes a spiral

search algorithm to locate tubes in the subsequent frame, facilitate jumping, skip over bad tube ends, and handle voids. Voids are places where there are no tubes because the tubesheet is supported by structure. To improve robustness, thresholds are established dynamically for light variations across the tubesheet. Brem and Nandhakumar divided the tube tracking process into three major stages:

1. correcting image data for the shape distortion caused by the camera viewing angle and hypothesizing and verifying the location of two tube-ends using a spiral search technique,

2. finding a better estimate of the location of the two tube-end centers on the image plane given initial estimates of their locations, and

3. displaying the row and column number of the tube closest to a predefined point in the video image.

Although their system has substantial merit, it also has several important limitations. The operator must manually compare the toolhead location to that computed by the robot control system. This is further complicated because the tubes being tracked rarely correspond to the tube that the robot toolhead is servicing. In addition, the set-up process is cumbersome and information displayed to the user is somewhat primitive. Finally the imaging geometry is modeled with only two vanishing points, rather than with the more general case of three vanishing points. The present work develops numerous improvements to refine and expand the capability of their methodology which include independent processing of interlaced image data fields; specialized edge operators; comprehensive 3-D mathematical imaging, tracking, and pose estimation models; integration with the robot control system; strategy for user interaction including the visual display of information; and scalability depending on the system requirements and available computational power.

Additional motion tracking work contained in the literature includes Jain et al. (1979), Jain (1984), Silven and Repo (1993), Aloimonos and Tsakiris (1991), Zheng and Chellappa (1995), Shapiro et al. (1995), and Lowe (1992). Major themes that are relevant to the present work consist of moving cameras rather than moving objects, passive rather than active tracking, monocular tracking in 3-D, tracking multiple features, processing image fields rather than frames, temporal feature correspondence, and real-time processing rates. Common research applications include road and vehicle following tasks, mobile robots, seam tracking, parts handling, automated assembly, particle velocimetry, and automatic surveillance systems. In the present work, camera mounting on the robot tool is arbitrary, but its motion is limited to three translations and one rotation. The tracking system must be capable of keeping track of the toolhead's 3-D position as the robot moves arbitrarily to various tube locations, without any direct control or knowledge of the robot's motions.

2.3 Robotics

Craig (1989), Sandler and Ben-Zion (1991), and Koivo (1989) discuss general robotic topics such as forward and inverse kinematics, path generation, dynamics, and control. In the context of the present work, the most useful portions of these references deal with spatial descriptions, transformation operators, robot calibration, and control.

Fu et al. (1987) discuss a wide variety of topics including various method of sensing such as range, proximity, tactile (touch), force, and torque. Low and high level vision is also introduced along with idea of rules-based intelligent systems. Finally, robot control methods are introduced.

Groover et al. (1986) provide a good overview of robot applications in manufacturing as well as many other fundamental robotic issues. The book divides applications into three main areas: 1) material transfer and machine loading/unloading, 2) processing

operations, and 3) assembly and inspection. The examples of particular interest include processes such as arc welding, spray coating, and assembly operations.

Mair (1988) offers a good view of the wide variety of electro-mechanical components found in modern robotics. Internal sensors such as resolvers, synchros, and optical shaft encoders are presented. External sensing includes photocells, lasers, vision, and acoustic systems. Many interesting applications are also discussed including arc welding, spray painting, and many different kinds of assembly operations.

Tidwell et al. (1991) provide an overview of the Cobra robot development. After discussing the background of the nuclear industry and steam generators, the conceptual design process that yielded the current manipulator configuration is presented. Mechanical design of the joints, links, tool coupling, and mezzanine are then briefly discussed. Robot calibration is divided into two parts. The first part is intrinsic calibration in which the link lengths and initial joint angles are measured. The second part is extrinsic calibration in which the position and the orientation of the robot base is found after installing the robot into the lower bowl portion of the steam generator. Finally, path planning, obstacle avoidance, and the user interface are discussed. This work is a predecessor to the current computer vision research and provides a good overview of the demanding requirements of steam generator robotics.

Shooter et al. (1992) introduce the idea of robot design for limited access workspaces (LAWs). A general approach is first presented, followed by the specific application of COBRA. In addition to providing a rigorous background, their work is also useful to understand the special constraints of steam generator environments.

Voruganti (1995) presents an analysis and design tool to improve the performance of robotic systems. Voruganti's method consists of identifying pose error sources, modeling the system, performing sensitivity analysis, and devising calibration procedures to achieve the desired performance. The calibration procedures consist of selecting a

suitable measurement system, collecting data, and optimizing for parameters that specifically describe the system. The method is applied to a specialized underwater mobile robot and a general-purpose Unimation PUMA robot. Voruganti minimizes end-effector pose errors by carefully characterizing the as-built geometry of the robot and its base pose. While this work is a significant advancement among calibration schemes, the control method is still representative of traditional indirect-feedback approaches. Both the work described in this dissertation and Voruganti's work have the same goal: to improve the positioning capability of robots. The two approaches proposed to achieve the common goal, however, are vastly different. These approaches are not mutually exclusive though, they can be combined to achieve the best of both control methods. The resulting robotic system would have a built-in vision measurement system to optimize the calibration parameters. The optimization could easily be performed periodically, particularly after the robot is serviced or re-configured. In addition, the vision system could provide pose information for small control corrections and quality assurance verification.

2.4 Controls

Many researchers have investigated using computer vision to augment the control of robots and tracking systems: Hashimoto et al. (1992), Herve et al. (1991), Kabuka et al. (1988), Papanikolopoulos and Khosla (1993), Westmore and Wilson (1991), and Corke (1993), among others. Although the approaches vary, they generally involve simple robotic systems or general systems with other types of sensor feedback. Some approaches utilize optical flow concepts and Kalman filters for position estimation. General Motors has also implemented vision to perform many practical tasks in automotive assembly including robotic welding (Schmidt, 1985), circuit board chip testing, and parts handling (Rossol, 1981). While the literature encompasses many methodologies and applications, it appears to lack the framing of a more general problem: controlling any robotic system using only visual feedback.

2.5 Mathematics and Statistics

Berger (1977) provides an excellent two-volume set in geometry which covers topics such as affine spaces, projective spaces, duality, and projective quadrics. Craig (1989) presents various spatial descriptions, mappings, operators, and transformation arithmetic that are commonly used in robotics. Riddle (1992) provides a general text for analytic geometry covering such topics as conic sections, planes in space, and symmetric line equations. Finally, Ott (1993) presents a general statistics text covering a wide variety of topics including simple, polynomial, and multiple linear regression.

Chapter 3

3. Hardware and Software

This chapter describes the overall hardware configuration, system components, and software architecture. The overall hardware configuration is first presented as it pertains to use in the field with steam generator servicing robots. Individual components such as cameras, lighting, image acquisition boards, processors, and display options are then described in more detail. Additional equipment that is used to develop and test the vision system, in both laboratory and full-scale mock-up settings, is also described. Finally, the software architecture is discussed, which consists of C code and functions from several different library sources running in a PC-DOS environment. Some additional hardware is presented in the next chapter because it is used specifically for system calibration and is better explained within that context.

3.1 Overall Hardware Configuration

Figure 3.1 illustrates the overall hardware configuration for the servicing and computer vision systems. The robot, servicing toolhead, and camera reside inside the steam generator, while their control hardware is mounted in portable enclosures in close proximity of the steam generator, inside the nuclear containment building. Two tethers,

consisting of more than 150 wires, connect the control enclosures with the equipment inside the generator. The robot and tool computer, vision computer, and system operator reside inside a control trailer. Video and control communication links run approximately 150 meters from the computers in the control trailer to the equipment inside the containment building.

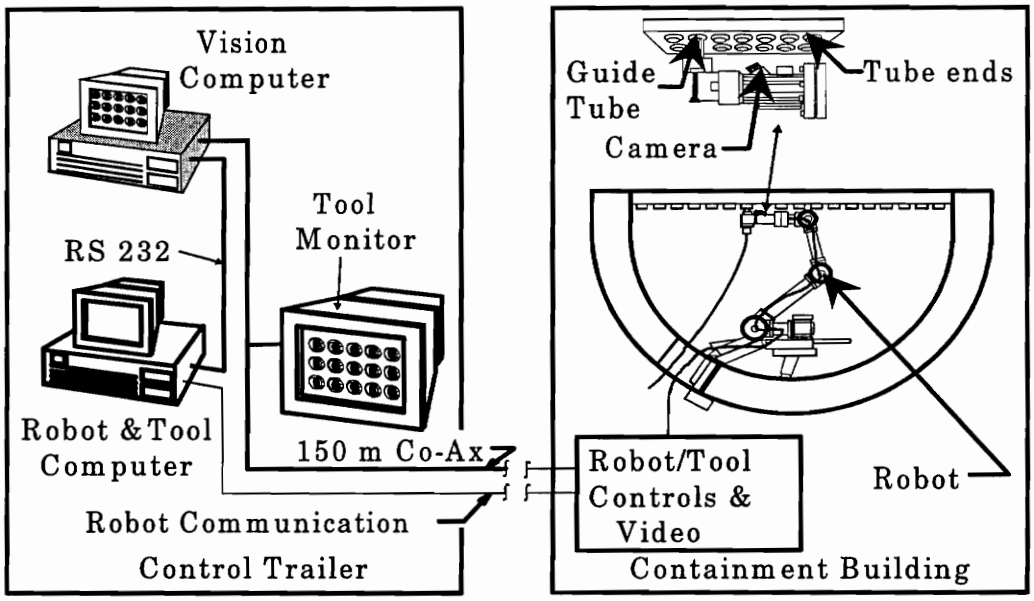


Figure 3.1: Overall Hardware Configuration

The primary additions to the existing robot servicing configuration are a vision computer and an auxiliary lighthead for the tool. The development effort was based on a personal computer with an image acquisition board. The analog video signal from the camera is sampled by the vision computer after it is displayed on the normal toolhead monitor. The system also includes an RS-232 serial link to communicate with the robot control computer. In future implementations, the separate vision computer may be eliminated and the system moved to the Hewlett Packard workstation that is used for robot control, tool control, and non-destructive evaluation acquisition tasks. In this implementation, however, the image acquisition board must also have on-board processors to perform all of the vision computations without additionally taxing the workstation.

3.2 Hardware Components

The vision system hardware is detailed in Figure 3.2. The toolhead mounted camera is a Sony model XC-75 with a 5 mm focal length lens (Sony Corporation of America, 1995). The camera acquires 30 images per second in interlaced form, meaning that the even image rows are generated in the first 1/60 of a second, followed by the odd image rows in the next 1/60 of a second. The images are transferred as an analog RS-170 signal to an ITI model VFG image acquisition board with twelve-bit frame memory (Image Technology Incorporated, 1995). The board digitizes the analog signal and stores the data as an array of 480 rows by 635 columns. Each element of the array is an eight-bit, gray-scale pixel representation, which corresponds to the intensity of light at that location in the image. The image is stored in the first eight bits, while the remaining four bits of memory can be used to display additional information on the video monitor along with the image. Although the original camera images are gray-scale, color can be added so that important information can be easily interpreted by the system operators.

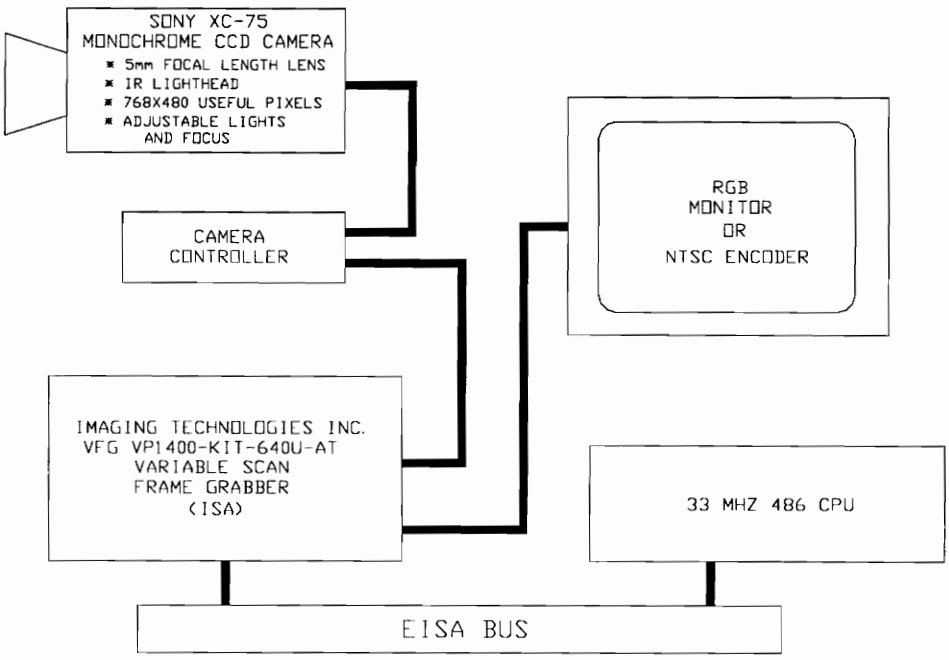


Figure 3.2: Computer Vision Hardware

The computations for processing image data, calibration, pose determination, and tracking are conducted on the host 33MHz Intel 80486 machine. The processed images and overlay information are then displayed on an RGB monitor or converted to the NTSC video standard by an encoder. The NTSC signals can then be integrated with the video switching system and displayed on the standard television monitors currently used with the servicing system in the field.

In addition to the computer vision hardware in the control trailer, lighting is also added to the toolheads inside the steam generator. A typical robot toolhead used for eddy current inspection of steam generator tubes is pictured under a tubesheet mock-up in Figure 3.3. The tool consists of a cylindrical, horizontal base and a vertical guidetube, which is used to pass eddy-current inspection probes up into the tubes. The camera (Figure 3.4) extends from the base on an angle to provide operators a view of the guidetube and several dozen tube-ends. Light is normally supplied by eight infrared, light emitting diodes (LEDs) mounted around the camera lens. The only modification to existing inspection tool and camera combinations is to add an auxiliary lighthead around the base of the guidetube.

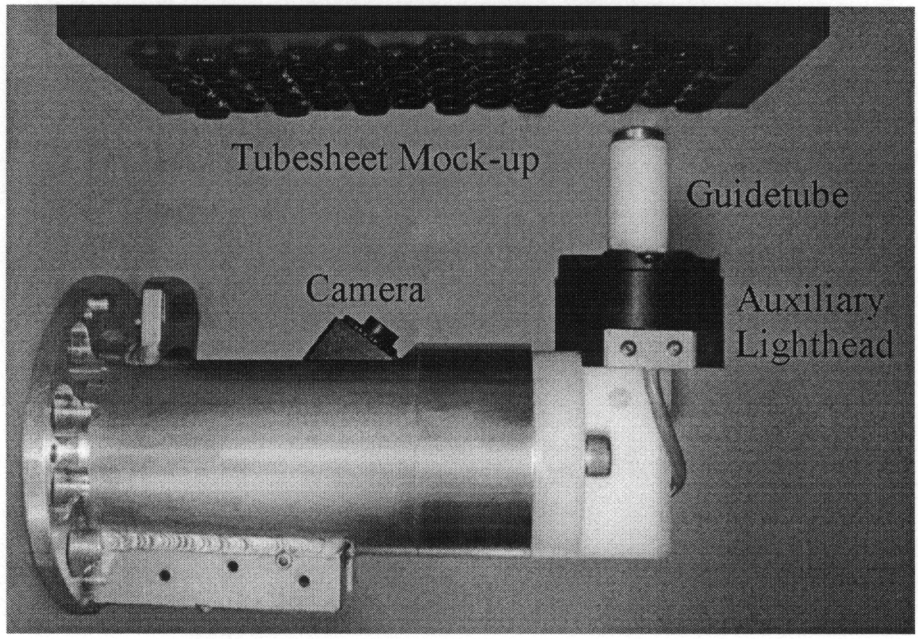


Figure 3.3: Robot Inspection Toolhead with Tubesheet Mock-Up Block

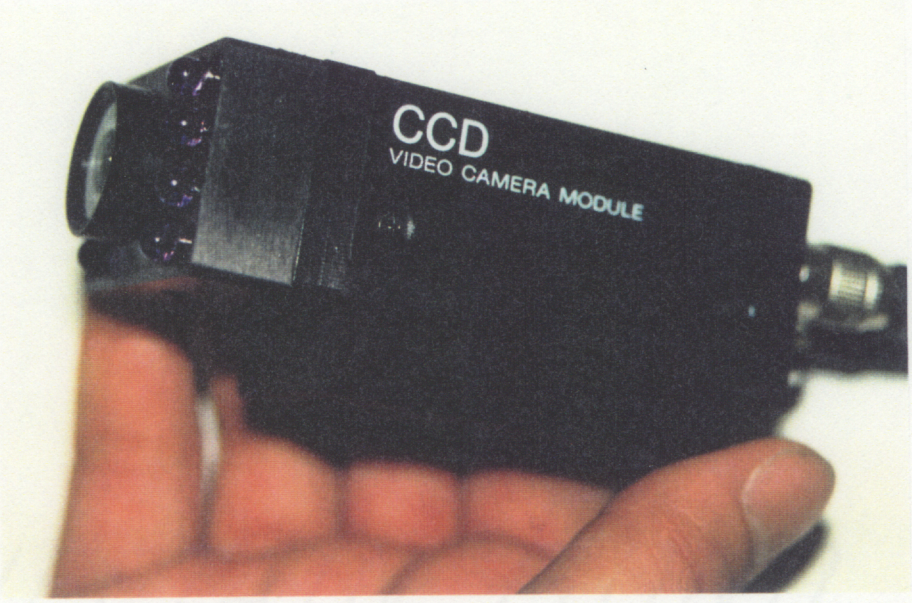


Figure 3.4: Sony CCD Array Camera with LED Light Ring

The lighthead consists of 12, small, incandescent light bulbs, as pictured in Figure 3.5. The light bulbs are controlled remotely with a variable 12 volt power supply and are directed straight up, toward the tubesheet. The lighthead is an important addition to the vision system because it produces high image contrast between the open tube-ends and the tubesheet surface.

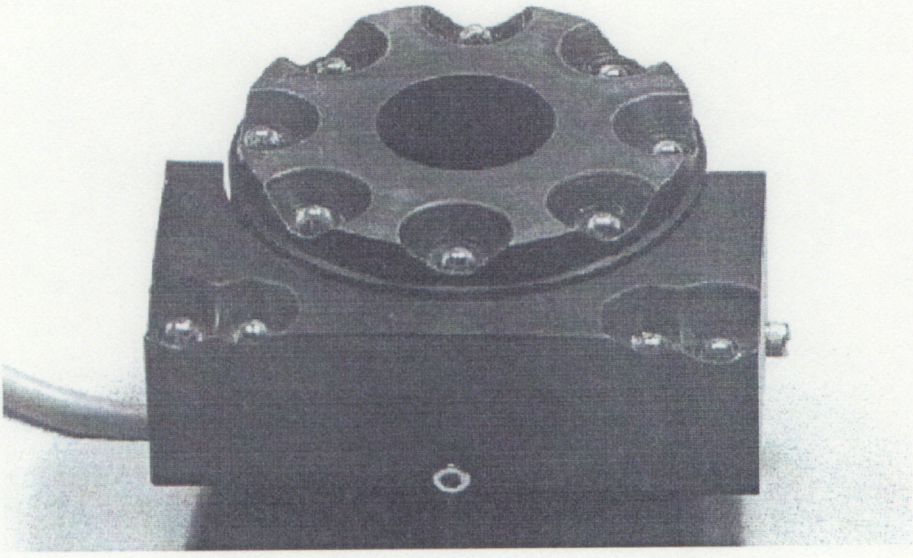


Figure 3.5: Robot Tool Auxiliary Lighthead

The advantages of the auxiliary lighthead over the camera-mounted LEDs are illustrated by the two toolhead camera views pictured in Figure 3.6 and Figure 3.7. Figure 3.6 shows the scene illuminated by the LEDs, while Figure 3.7 shows the same view illuminated by the auxiliary lighthead.

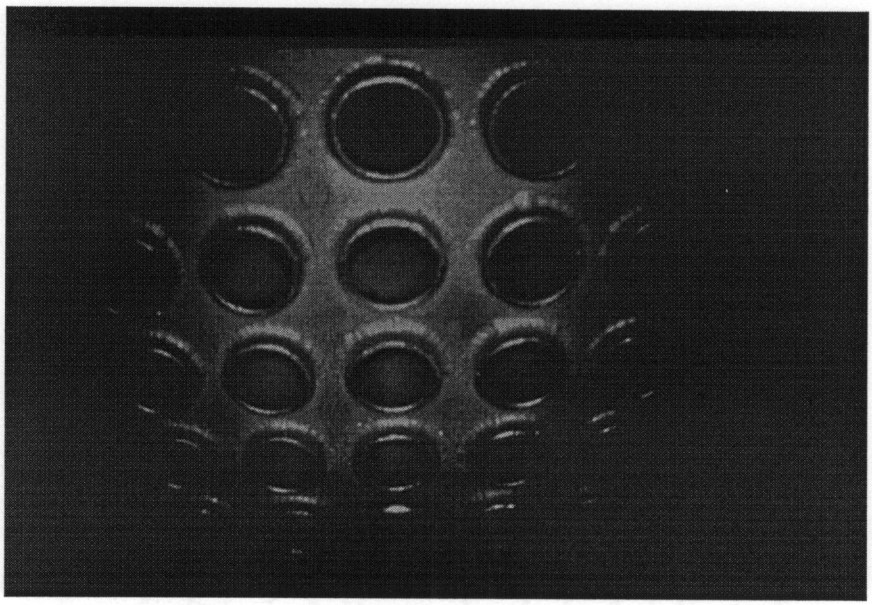


Figure 3.6: Toolhead Camera View without Lighthead

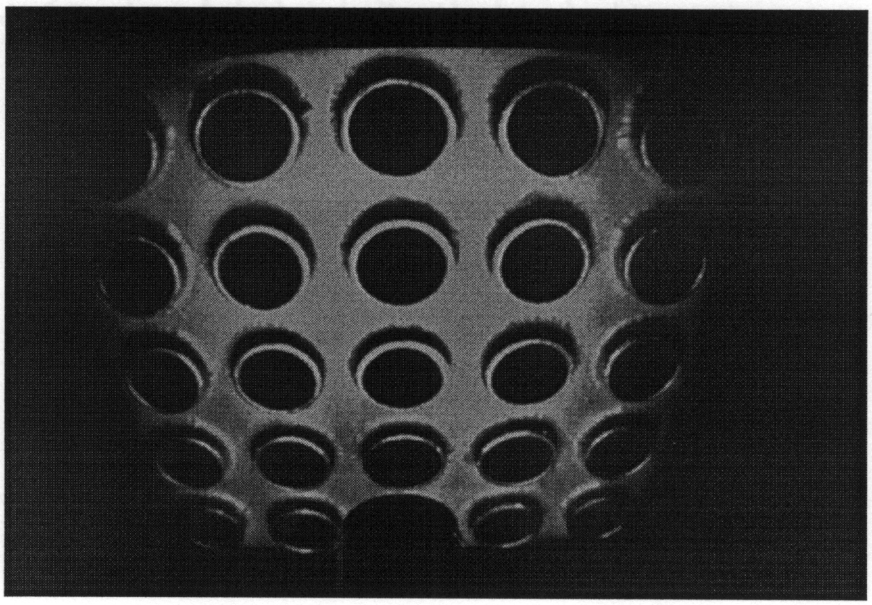


Figure 3.7: Toolhead Camera View with Lighthead

Light levels in the image are generally low when the scene is illuminated only by the camera-mounted LEDs. Additionally, since the light originates from the viewing direction, which is angled to the tubesheet, the inside portions of many of the tubes are illuminated and the tube-end edges are indistinct. When the auxiliary lighthead is used instead of the LEDs, the image quality improves dramatically. Note how the tube interiors are uniformly dark, the tubesheet surface is bright, and the tube-end edges are distinct. The edges are also more accurately represented in the image since the light source is directed normal to the surface of the tube-ends. The auxiliary lighting facilitates image processing and enables the vision system to meet both the speed requirements for tracking and the accuracy requirements for pose determination. Lighting is particularly important when low-cost equipment is to be used for real-time applications.

3.3 Laboratory Experimental Arrangement

To initially develop and test computer vision algorithms, the experimental configuration shown in Figure 3.8 was constructed in the laboratory.

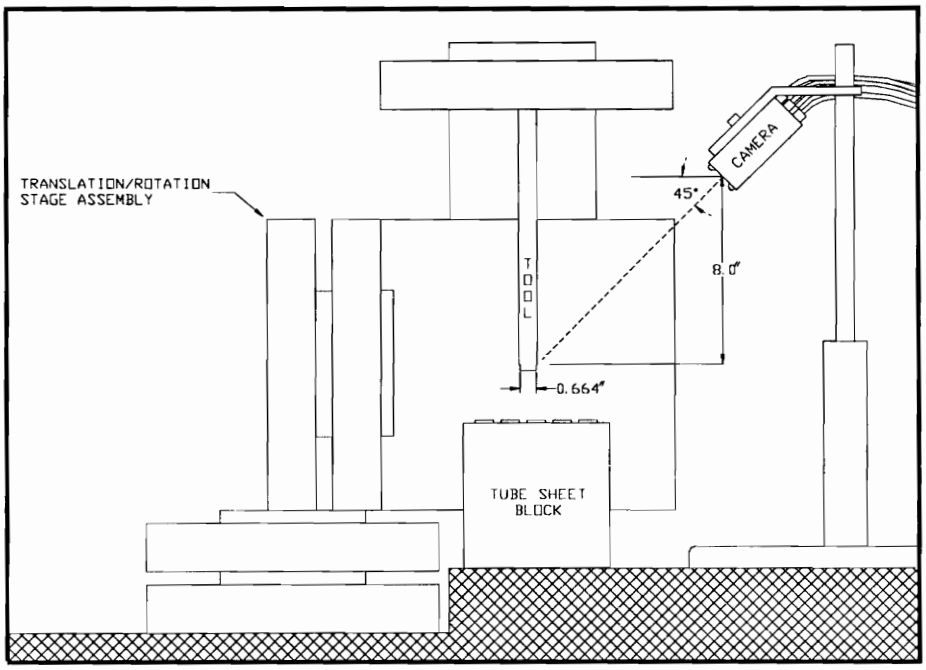


Figure 3.8: Laboratory Experimental Configuration

The laboratory equipment includes the computer vision hardware already described along with a metrology table, six-axes pose measuring system, light isolation chamber (not shown), tubesheet mock-up block, and various adapter brackets. The measurement system is securely attached to the metrology table and consists of three translation and three rotation stages. The translation stages are accurate to 0.005 mm (0.0002 in.), while the rotation stages are accurate to 5 minutes-of-arc (Ealing Electro-Optics, 1995). The isolation chamber encloses the metrology table and prevents background light from interfering with the simulated environment. Various adapter brackets were fabricated so that either a guidetube (shown in the figure) or an entire toolhead could be mounted to the axes system and moved over the mock-up block. The experimental laboratory configuration was used to investigate lighting as well as to characterize the accuracy of the vision algorithms.

3.4 Full-Scale Mock-Up

A full-scale steam generator mock-up was used to test tracking and non-contacting manipulator base calibration algorithms, with a variety of robots. Figure 3.9 shows a servicing robot called ROGER installed inside the mock-up. Figure 3.9 also shows the numerous electrical, pneumatic, and hydraulic lines that are used to control the robot and servicing toolheads. The complex array of cabling underscores the importance of developing a computer vision system that requires minimal additional hardware and cabling inside the steam generator. This is also the fundamental reason why developing monocular vision for use with existing toolhead cameras has inherent advantages over stereo vision, acoustic, or laser systems. The steam generator mock-up normally consists of quarter-sphere inlet and outlet plenums, where the robots are installed to service the generator, and a drilled plate that represents the tube bundle. While this mock-up is an excellent mechanical simulation, it must be modified so that it also visually simulates the color and reflectance characteristics of the oxidized surfaces inside the real generators. Figure 3.10 shows a toolhead camera view of the mock-up, where tubes have been installed in the drilled holes and the surfaces painted to simulate actual plant conditions.

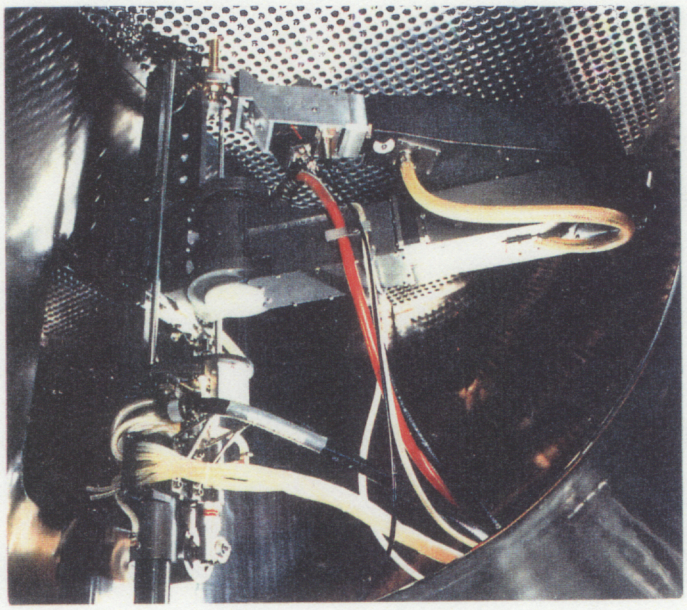


Figure 3.9: ROGER in a Steam Generator Mock-Up

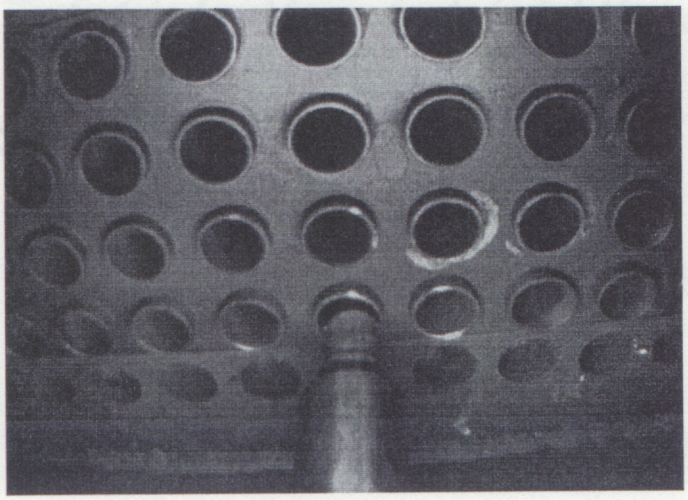


Figure 3.10: Toolhead Camera View of Mock-Up

In addition to carefully simulated steam generator mock-ups, video tapes of actual steam generators were also used to test the vision system. Video tapes were made by installing auxiliary lighting on the robot tools and recording toolhead video as the robot moved under the tubesheet. The simulated laboratory and mock-up environments along with the video tape provided excellent test beds prior to the costly site tests, where the vision system was deployed at various nuclear power stations.

3.5 Software Architecture

The software architecture consists of C code (Kerningham and Ritchie, 1988) running in a PC-DOS environment. The code consists of specialized algorithms, along with high and low-level library functions for the image acquisition board. The overall software is divided into five programs as follows:

1. internal system calibration image processing module,
2. internal system calibration computational module,
3. external system calibration module,
4. pose determination module, and
5. tracking module.

Further details on these modules are explained in subsequent chapters, along with how they are integrated into the overall vision and robot control systems.

Chapter 4

4. System Calibration

To obtain high accuracy with computer vision, the imaging system must be accurately modeled and calibrated. This chapter presents a conceptual imaging model that completely describes the relationship between objects in space and their image data within the computer. Mathematical models are then presented for ideal and real imagers. The models are used to calibrate both internal and external system parameters. Specially developed hardware and a well-known, popular algorithm are first used to solve for internal system parameters. Geometric knowledge of the working environment is then used to solve for camera external pose parameters as well as an internal parameter subset.

4.1 Conceptual Imaging Model

The mapping of information in a typical computer vision system is illustrated in Figure 4.1. Objects in space are imaged by a lens system onto an imaging sensor, such as a CCD array. The CCD electronics quantizes the image and produces a clocked analog signal that is transferred to a digitizer board in a computer. The digitizer converts the analog signal into digital values corresponding to light intensity, which are then stored as an array in computer memory.

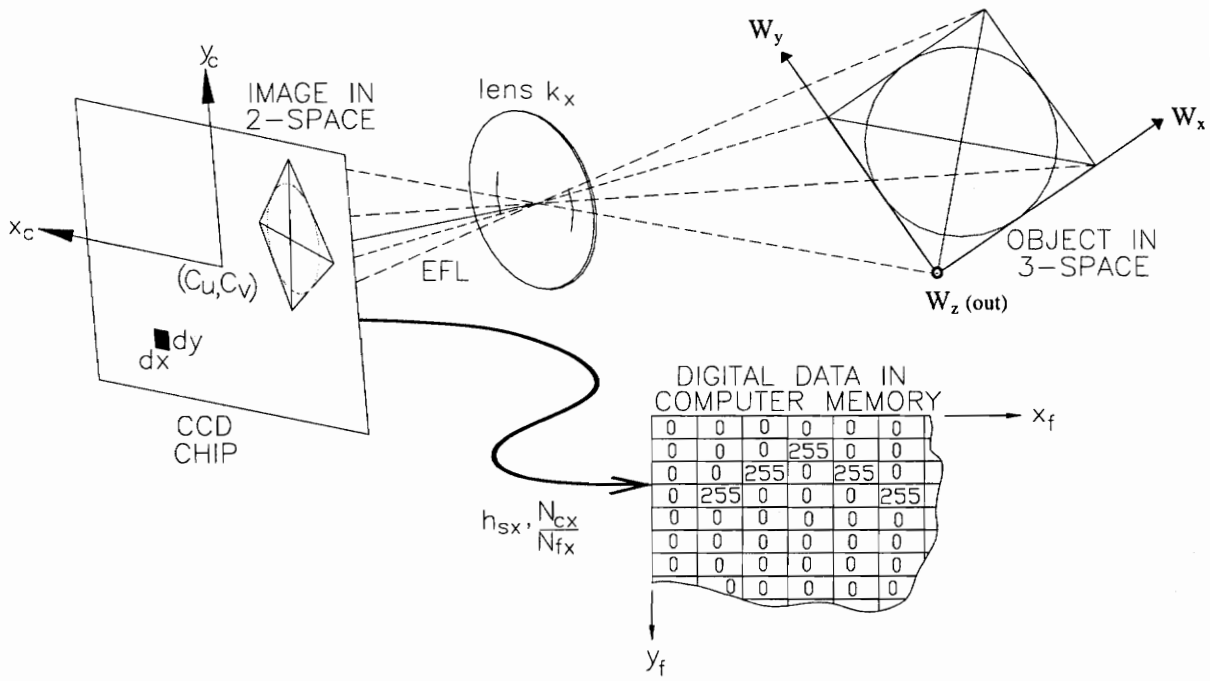


Figure 4.1: Conceptual Imaging Model for a Computer Vision System

4.2 Ideal Imaging System

An ideal relationship between a point in space (W_x, W_y, W_z) and its image (x_c, y_c) is given by Equation (4.1). The homogeneous representation of point **W** is first pre-multiplied by a general spatial transform which describes the relationship between the object coordinate frame, where **W** is defined, and the camera coordinate frame (Craig, 1989). Once **W** is expressed in the camera frame, it can be imaged by pre-multiplying a perspective transformation.

$$\begin{bmatrix} h\cdot x_c \\ h\cdot y_c \\ h \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{-1}{EFL} & 1 \end{bmatrix} \cdot \begin{bmatrix} c\alpha\cdot c\beta & c\alpha\cdot s\beta\cdot s\gamma - s\alpha\cdot c\gamma & c\alpha\cdot s\beta\cdot c\gamma + s\alpha s\gamma & x \\ s\alpha\cdot c\beta & s\alpha\cdot s\beta\cdot s\gamma + c\alpha\cdot c\gamma & s\alpha\cdot s\beta\cdot c\gamma - c\alpha\cdot s\gamma & y \\ -s\beta & c\beta\cdot s\gamma & c\beta\cdot c\gamma & z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} W_x \\ W_y \\ W_z \\ 1 \end{bmatrix} \quad (4.1)$$

where, $z_c = 0$ *at the image plane,*

EFL is the distance along the optical axis (z_c) *between the optical point of projection within the lens system and the image plane,*

h is a homogeneous coordinate,

c and s are shorthand forms of cosine and sine respectively, and

rotations α, β, γ *and translations x, y, z are with respect to a fixed world coordinate system.*

The ideal image coordinates are then recovered by dividing the first and second elements of the left-hand-side of Equation (4.1) by the third element, h.

The perspective imaging transformation of Equation (4.1) can be derived by referring to the simple case illustrated in Figure 4.2. In this example, point P is already defined in the camera coordinate system so it can be imaged through the optical projection point without first applying the spatial transform.

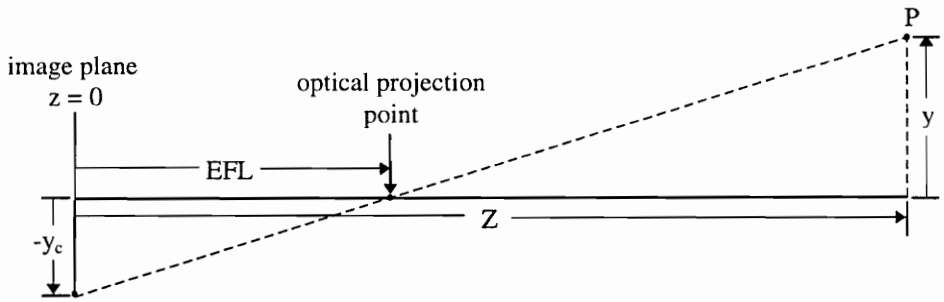


Figure 4.2: Deriving the Perspective Imaging Transform

Observing the similar triangle geometry of Figure 4.2, the ratio given in Equation (4.2) can be written and rearranged to form Equations (4.3) and (4.4), respectively.

$$\frac{-y_c}{EFL} = \frac{y}{z - EFL} \tag{4.2}$$

$$y_c = \frac{-EFL \cdot y}{z - EFL} \tag{4.3}$$

$$y_c = \frac{y}{1 - \frac{z}{EFL}} \tag{4.4}$$

Similar results can also be generated for the x_c coordinate and then used to construct the appropriate perspective transformation as shown in Equation (4.1).

4.3 Modeling Real Imaging Systems

Equation (4.1) is an ideal relationship because it uses an ideal pinhole instead of an actual lens system. Additional relationships are needed to model real lenses, imaging sensors, and digitizer boards. Referring again to Figure 4.1, a lens distortion coefficient (k_X) is used in a radial, parabolic model to correct pin cushion or barrel distortion. The image center (c_U,c_V) is the location where the optical axis intersects the CCD image plane. The horizontal scale factor (h_{SX}) together with the number of horizontal camera (N_{cx}) pixels and digitizer frame memory pixels (N_{fx}) relate the width of camera and frame pixel elements. A vertical scale factor is not required, since the number pixel rows for the camera and frame buffer are forced to match exactly by the RS-170 video standard. The width and height of pixel elements on the CCD array are given by dx and dy respectively.

Lens distortion results from the interaction of field curvature, flat image planes, and aperture stops. Field curvature is best explained with geometrical optics. If light rays enter the lens system at a significant angle with the optical axes, the light rays converge on a curved surface as shown in Figure 4.3 (Longhurst, 1967). The light rays then diverge and image as a blur on the flat image plane. If aperture stops are used in the optical system, some of the light rays are blocked and the resulting image is sharper, but distorted. A stop at location "B" produces barrel distortion, while a stop at location "P"

produces pin cushion distortion, as illustrated in Figure 4.4. Lens distortion can be quite extreme when objects are imaged at close range with wide angle lens systems.

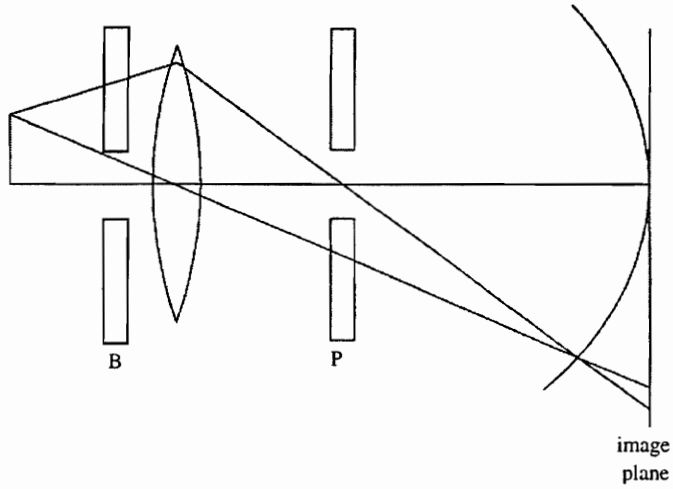


Figure 4.3: Lens Distortion with Aperture Stops

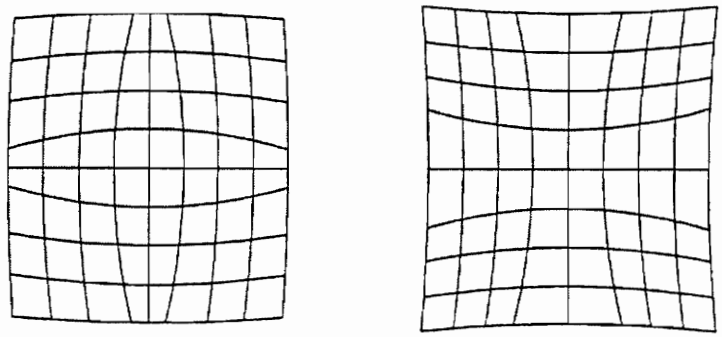


Figure 4.4: Barrel and Pin Cushion Distortion

The extent of the barrel distortion for the wide angle lenses such as those used in steam generator robotics is evident in Figure 3.7. The straight rows and columns of tubes near the outer edges of the image appear to curve inward.

Coordinates in the computer's frame buffer (x_f , y_f) are related to the corresponding ideal pin-hole camera coordinates (x_c , y_c) with Equations (4.5) through (4.9) (Tsai, 1987):

$$x_d = -\frac{dx}{h_{sx}}\left(\frac{N_{cx}}{N_{fx}}\right)(x_f - c_u) \tag{4.5}$$

$$y_d = -dy(y_f - c_v) \tag{4.6}$$

$$r_d^2 = x_d^2 + y_d^2 \tag{4.7}$$

$$x_c = x_d(1 + k_x r_d^2) \tag{4.8}$$

$$y_c = y_d(1 + k_x r_d^2) \tag{4.9}$$

where, x_d *and* y_d *are intermediate values,*

x_f *and* y_f *are in units of frame buffer pixels in computer memory, and*

x_c *and* y_c *are in the same units as dx and dy (usually mm).*

After appropriate calibrations, Equations (4.1) and (4.5) through (4.9) will be used to accomplish the following tasks:

1. Map the locations of point data in computer memory to corresponding point locations on the ideal image plane. The points on the ideal image plane correspond to the object being imaged by an ideal pin-hole camera. Since the complexities related to the lens system, CCD sensor, and digitizer have been accounted for using Equations (4.5) through (4.9), ideal points can be readily used with geometric relationships.

2. Predict the locations of point data in computer memory that correspond to point data on the ideal image plane or in object space.

3. Determine the spatial transform that describes the relationship between the object and camera coordinate frames and, in some cases, also determine the perspective transformation.

4.4 Calibrating System Parameters

Calibrating computer vision systems involves determining the values of both internal and external parameters. Internal parameters consist of those used in Equations (4.5) through (4.9) and the effective focal length (EFL) used in the perspective transformation of Equation (4.1). Camera parameters (dx, dy, and N_{cx}) vary depending on the camera in use and are generally available from the manufacturer. Similarly, N_{fx} is available from the digitizer board manufacturer or can easily be measured in computer frame memory. The remaining parameters (EFL, (c_u , c_v), h_{sx}, and k_x) must be calibrated for particular camera, lens, and digitizer combinations. The external parameters (x, y, z, α, β, γ) describe the relationship between the camera and robot tool coordinate frames and are embedded in the spatial transform of Equation (4.1). The external pose parameters must be calibrated after the camera is installed in the toolhead or whenever the relationship between the guidetube and camera is adjusted.

4.4.1 Internal Parameters

Tsai (1987) and Lenz and Tsai (1988) propose a two-stage technique that determines the internal parameters EFL, (c_u , c_v), h_{sx}, and k_x, thereby enabling off-the-shelf cameras to be used for accurate geometric measurements. Camera lens distortions can be divided into two components, radial and tangential. For many cameras, the tangential portion is negligible, and the distortion can be treated as radial only. Their method simplifies the large scale nonlinear problem with a radial alignment constraint. The internal camera/lens/digitizer parameters are calibrated using the Lenz and Tsai method and the specially designed hardware pictured in Figure 4.5.

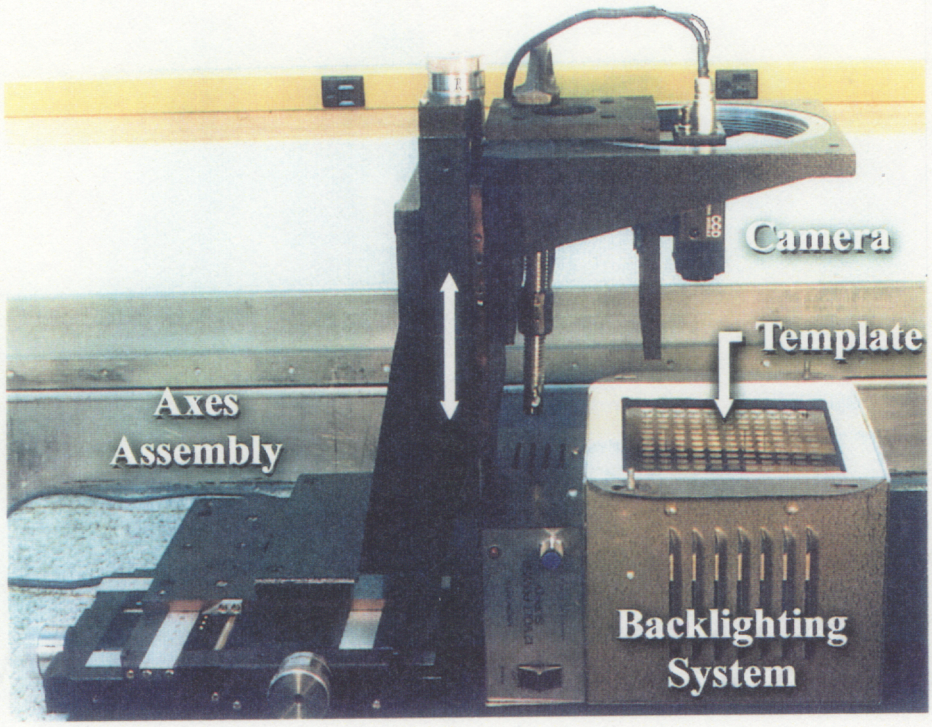


Figure 4.5: Vision System Calibration

The hardware consists of a vertical translation stage, precision template, back-lighting system, and the vision system hardware. The translation stage has a vertical travel of 100 mm and is accurate to 0.005 mm (Ealing Electro-Optics, 1995). The template consists of an array of 100 transparent circles formed by selectively applying a low-reflective, chrome coating to a 1.5 mm thick, soda-lime, glass plate. The template is fabricated with a photographic process which yields a circle center-to-center accuracy of 2.5 microns (GAGE-Line Technologies, 1995). The back-lighting system consists of an array of small incandescent light bulbs and a glass diffusion plate. The bulbs are the same type as those used for auxiliary tool lighting. The light array is directed downward so the light must first reflect off the box interior surfaces before passing through the diffusion glass and back-lighting the template. This arrangement produces uniform lighting which facilitates thresholding and region-based image processing.

In the calibration, images of the array are acquired by the digitizer board at two camera heights perpendicular to the template. The centroids of all the dots in computer frame memory at each height, the precise distance between the two camera heights, and the center spacing of the array dots on the template are then used in a two-stage optimization process to solve for the internal parameters (Tsai, 1987 and Lenz and Tsai, 1988). The software for the internal parameter calibration is divided into two modules. The first module performs the image processing and organizes the data in several files for the second module which then performs the optimization.

For highly accurate pose determination, such as that required for manipulator base calibration, this calibration should be performed once in the laboratory for each camera/lens/digitizer combination and focus setting. For the steam generator robotics work, once the focus has been set in the laboratory, it does not normally need to be adjusted because the aperture stop produces a large depth-of-field. If the vision system will only be used for tracking and position correction then this calibration may only be performed once for each type of camera/lens/digitizer combination. The effective focal length for a particular camera could then be determined as part of the external parameter calibration.

4.4.2 External Parameters

The external parameters (x, y, z, α, β, γ) describe the relationship between the camera and robot tool coordinate frames and are embedded in the spatial transform of Equation (4.1). Since toolhead and camera mounts are adjustable, the external parameters must be calibrated during steam generator servicing operations. The calibration is performed after the camera is installed in the toolhead and the toolhead is loaded on the robot, inside the steam generator. The goal of the external calibration is to compute the orientation of the guidetube axis and the position of the guidetube tip in the camera coordinate system. This pose information will then be used later to compute the guidetube-to-tubesheet transform and which tube is directly above the guidetube.

The calibration process is summarized here, with further details provided when pose determination is discussed in Chapter 5:

1. First, the robot is moved so that the tool is docked onto a tube-end inside the steam generator.

2. The toolhead camera image is processed and the locations of specific image features extracted.

3. The feature data is mapped to the ideal image plane using Equations (4.5) through (4.9).

4. The feature locations on the ideal image plane are used as the (x_c , y_c) points in Equation (4.1).

5. The known tubesheet geometry is used to define a local coordinate system and to generate object data corresponding to the image features. The object data is then used as the (W_x , W_y , W_z) points in Equation (4.1).

6. Using at least four points (12 equations), an objective function is constructed and minimized to obtain the values of $(x, y, z, \alpha, \beta, \gamma)$ which give the best match between the image and object data.

The effective focal length may be known from the internal system calibration or it can also be included in the minimization. Processing tubesheet information gives the relationship between the camera and the tool because the tool's guidetube is docked, or coupled, to a tube-end.

Chapter 5

5. Pose Determination

The task of pose determination is to accurately estimate the three position and three orientation parameters that describe the relationship between the camera and object coordinate systems. As discussed earlier, there are a number of methods that can be employed. This chapter focuses on single-view, monocular imaging, where precise geometric knowledge of objects can be used with a single view from a single camera. Image processing is first discussed, including edge detection, feature-based regression, and point-wise correspondence between object and image spaces. Various strategies are then discussed to determine pose, including perspective projection of three, four, and N-points. At least three points are required to solve for the pose parameters, but the three-point formulation yields eight solutions. Several formulations exist which yield a unique, closed-form solution for four points, but they are limited to planar points which form rectangles or quadrangles. The most general and accurate results are shown to be obtained with an N-point optimization method.

Pose determination is then applied to calibrate camera parameters (as discussed in Chapter 4), calibrate the robot base position, verify the accuracy of the tool pose, and provide a correction to support difficult tool-to-tube insertion tasks. Results from

simulation, sensitivity analysis, and field testing with the actual robot control system quantify the performance of the vision system and show that its design specifications have been met.

5.1 Image Processing

To obtain high accuracy with single-view, monocular imaging, the image is processed after the robot toolhead has come to rest under the tubesheet. Since the camera is at rest, temporal frame averaging can be used in the presence of noisy images. As an option, the system first averages 32 frames in about one second. Tube-end features are then selected manually by the operator or automatically by the tracking system (described in Chapter 6). Manual selection is used for initially calibrating the camera and robot base positions, while automatic selection is used for rapid pose verification. In the manual case, the system operator positions a cross-hair in the darkened region of at least four tube-ends and enters their row/column numbers. The row/column numbers are used with the known tubesheet geometry to build an accurate, geometric, object model from the arbitrarily selected features. After the features are selected and identified, edge data can be detected in the image and regressed to define feature curves.

5.1.1 Edge Detection

Edge detection is simplified by using the starting points within the darkened tube-ends. The auxiliary tool lighting helps to create a high contrast image with sharply defined features. First, the image is searched vertically from the starting point with two bands as shown in Figure 5.1. After the top and bottom edge boundaries are found, their locations are bisected and the image is search horizontally. The left and right edge boundaries are then bisected to yield the tube-end center. A second vertical search is performed from the center to find the maximum top and bottom edge boundaries and to adjust the center location. The center is used as a starting point for an edge operator that searches for edges within the space defined by the boundary locations.

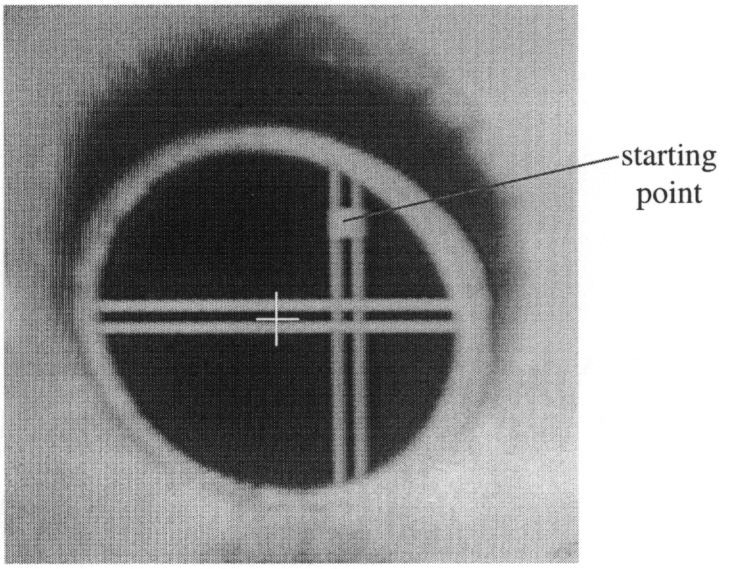


Figure 5.1: Searching for Tube-End Boundaries

Each search for edge data consists of two parallel paths which are one pixel wide and separated by one pixel. The separation distance improves the robustness of the search in the presence of image noise. Edges are found by applying a gradient edge operator to each pixel on the two search paths. The gradient results for corresponding pixels on the two paths are then added together to provide a measure of the confidence for the existence and location of the edge point. The gradient operator, defined in Equation 5.1, uses the intensity values of 24 neighboring pixels on the two paths.

$$\begin{aligned} G_{(u,v)} = {} & I_{(u+1,v+2)} + I_{(u+2,v+2)} + I_{(u+3,v+2)} + I_{(u+4,v+2)} \\ & + I_{(u+5,v+2)} + I_{(u+6,v+2)} - I_{(u-1,v+2)} - I_{(u-2,v+2)} \\ & - I_{(u-3,v+2)} - I_{(u-4,v+2)} - I_{(u-5,v+2)} - I_{(u-6,v+2)} \\ & + I_{(u+1,v-2)} + I_{(u+2,v-2)} + I_{(u+3,v-2)} + I_{(u+4,v-2)} \\ & + I_{(u+5,v-2)} + I_{(u+6,v-2)} - I_{(u-1,v-2)} - I_{(u-2,v-2)} \\ & - I_{(u-3,v-2)} - I_{(u-4,v-2)} - I_{(u-5,v-2)} - I_{(u-6,v-2)} \end{aligned} \tag{5.1}$$

$I_{(u,v)}$ in Equation (5.1) represents the pixel intensity value at column u and row v in the image array, while $G_{(u,v)}$ is the computed gradient value. Figure 5.2 shows a typical tube-end along with the corresponding pixel intensity and gradient profiles as calculated with Equation (5.1). Note that the gradient values are actually ten times those indicated on the graph and that the gradients are computed from the center estimate outward. Pixel

intensity gray-level values range between 0, which corresponds to black, and 255, which corresponds to white. The computations are held to a minimum by terminating the search when the gradient value exceeds a dynamically established threshold. The gradient operator produces a large value when an edge is encountered, yet it is relatively insensitive to image noise. The operator achieves this important combination by investigating many neighboring pixels on two separated paths that extend both far ahead and behind the pixel being investigated. If an edge that is difficult to detect is encountered, the dynamic thresholds are reduced, and the search is conducted once again.

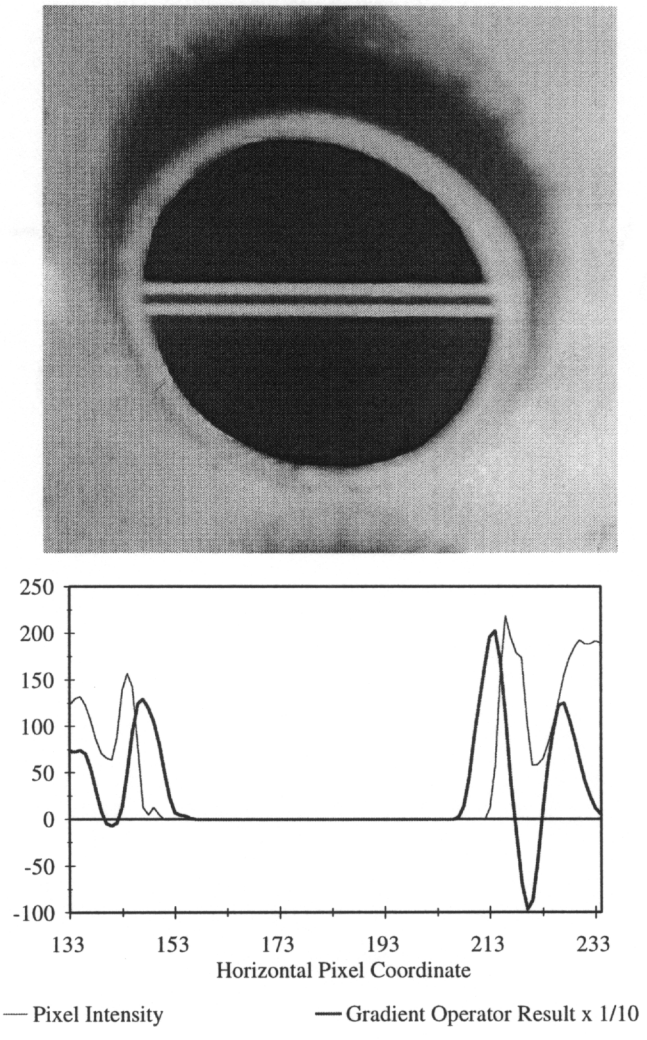


Figure 5.2: Horizontal Intensity and Gradient Profiles for a Typical Tube-End

Figure 5.3 shows how the gradient operator is used to scan the image and detect tube-end edges. First the horizontal image rows are searched outward from the vertical tube-end center line. The rows are indexed away from the center, moving up and down in the image until the search boundaries are reached. Since edges are best detected when they are orthogonal to the search direction, the scanning is terminated shortly before the top and bottom of the tube-end is reached. Next, the vertical image columns are searched outward from the horizontal tube-end center line. The columns are also indexed away from the center, moving left and right in the image until the search boundaries are reached. During the search, a continuity constraint is enforced to eliminate spurious edge data. The constraint requires that adjacent edge data be spatially consistent, but does not require connectivity so some weak edge data can be skipped. The dynamic thresholds for the edge operator are set at 75% of the previous gradient value and reduced by 50% if the edge is difficult to detect.

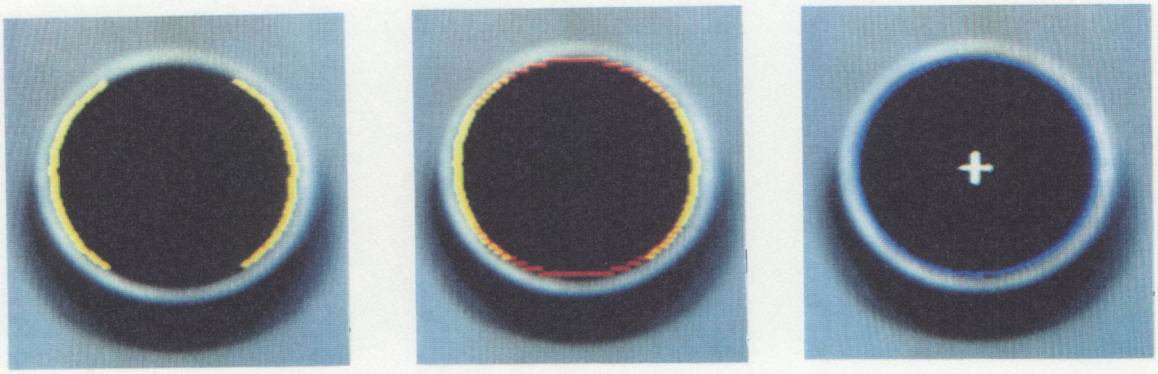


Figure 5.3: Locating Tube-End Edges by Scanning the Image with the Gradient Operator

The horizontal and vertical searching technique is well suited for the vision hardware, which is more efficient at accessing vertical and horizontal lines in the image than individual pixel values. For this reason, horizontal and vertical searching is used instead of a radial technique, even though a radial search would provide slightly more orthogonality between the edge and the search path. The gradient edge operator is also well suited for the vision hardware and also takes advantage of the object geometry and scene illumination.

Auxiliary tool lighting facilitates edge detection by intensifying the edges of interest and removing false edges. Figure 5.4 shows a picture of the tubesheet with boron deposits that typically form as a result of steam generator water chemistry. These deposits create false edges even with fairly good lighting inside the steam generator. Figure 5.5 shows the tubesheet when illuminated by the auxiliary tool lighthead (this is a slightly different view than Figure 5.4). The edges, corresponding to the tube-end inside diameter, are distinct and the false edges are successfully removed.

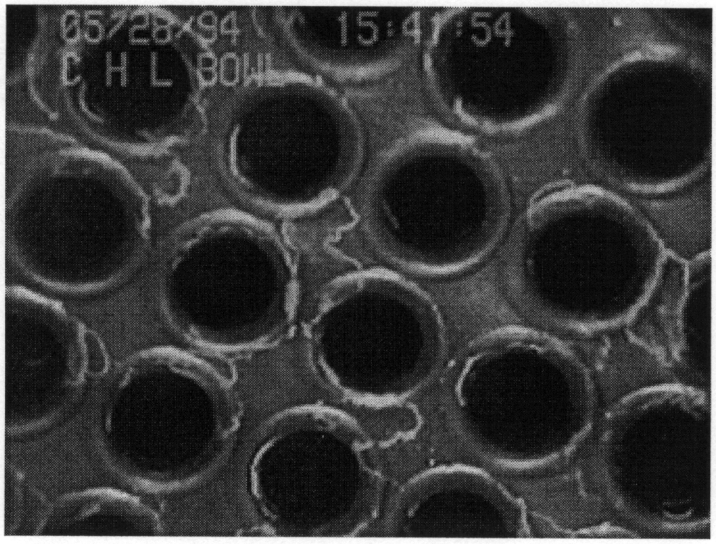


Figure 5.4: False Edges Created by Boron Deposits

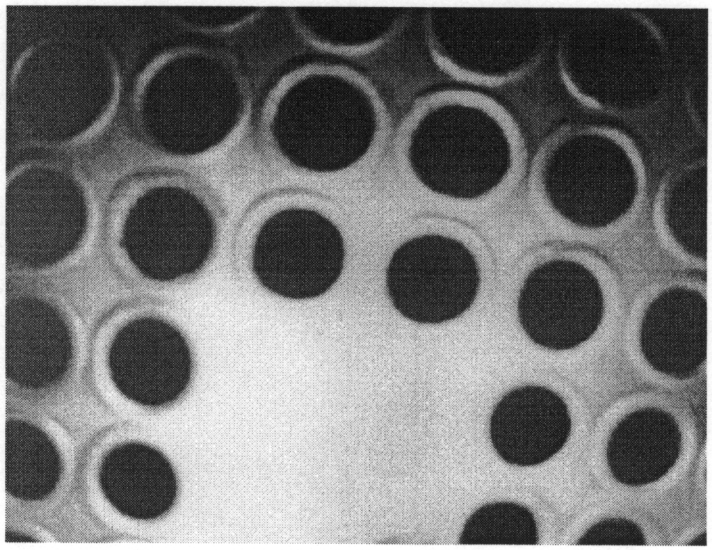


Figure 5.5: Auxiliary Lighting Removes False Edges

5.1.2 Feature-Based Regression

After the edge point data has been extracted from the image, it is corrected using equations (4.5) through (4.9) and regressed to the appropriate image feature curve. In general, the geometry of an optical imaging system with a circular object feature is an oblique circular cone intersected by a plane. Apollonius of Perga proved more than 2000 years ago that this geometry produces a conic section (Boyer, 1968). For the case of imaging, the conic formed on the image plane is an ellipse. A proof is provided here based on Mundy and Zisserman (1992) and Berger (1977):

The general form of a conic in the Euclidean plane is given by Equation (5.2):

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{5.2}$$

Conics can also be expressed in the projective plane using homogeneous coordinates as defined in Equation (5.3). Substituting Equation (5.3) into Equation (5.2) yields Equation (5.4), which can then be multiplied by z'^2 to obtain the quadratic form given in Equation (5.5). This can also be expressed using the weighted inner product given in Equation (5.6), where the weighting matrix C is symmetric for a conic.

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \quad \Rightarrow \quad x = \frac{x'}{z'} \qquad y = \frac{y'}{z'} \tag{5.3}$$

$$a\left(\frac{x'}{z'}\right)^2 + b\frac{x'}{z'}\frac{y'}{z'} + c\left(\frac{y'}{z'}\right)^2 + d\frac{x'}{z'} + e\frac{y'}{z'} + f = 0 \tag{5.4}$$

$$ax'^2 + bx'y' + cy'^2 + dx'z' + ey'z' + fz'^2 = 0 \tag{5.5}$$

$$\begin{pmatrix} x' & y' & z' \end{pmatrix} \begin{bmatrix} a & \frac{b}{2} & \frac{d}{2} \\ \frac{b}{2} & c & \frac{e}{2} \\ \frac{d}{2} & \frac{e}{2} & f \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = 0 \quad \text{or} \quad P^t CP = 0 \tag{5.6}$$

Applying a projective transformation T to a point P in Equation (5.7) gives a relationship that can be used to apply a projective transformation to the conic as shown in Equation (5.8). The transformation T is a general projective mapping which includes perspective viewing as a subset.

$$P' = TP \qquad \Rightarrow \qquad T^{-1}P' = P \tag{5.7}$$

$$\left[T^{-1}P'\right]^t C\left[T^{-1}P'\right] = 0 \tag{5.8}$$

Using the matrix property of Equation (5.9), Equation (5.8) can then be expressed as the weighted inner product given in Equation (5.10).

$$(AB)^t = B^t A^t \tag{5.9}$$

$$P'^t\left[T^{-1}\right]^t C\left[T^{-1}\right]P' = 0 \qquad \text{or} \qquad P'^t C' P' = 0 \tag{5.10}$$

To show that the projection of a conic yields a conic, C' must be shown to be symmetric, which implies showing the following:

$$C'^t = C' \qquad \text{or} \qquad \left\{\left[T^{-1}\right]^t C\left[T^{-1}\right]\right\}^t = \left[T^{-1}\right]^t C\left[T^{-1}\right] \tag{5.11}$$

Using the matrix properties in Equations (5.9) and (5.12) and remembering that by definition $C^t = C$, showing that C' equals its transpose is easily accomplished in Equation (5.13).

$$\left(A^t\right)^t = A \tag{5.12}$$

$$\left[T^{-1}\right]^t C^t\left\{\left[T^{-1}\right]^t\right\}^t = \left[T^{-1}\right]^t C\left[T^{-1}\right] \tag{5.13}$$

Since the projection of a circular object feature produces an ellipse on the camera's image plane, the edge point-data will be regressed to an ellipse image feature.

Beginning again with Equation (5.2), we can use the general form of the conic to formulate an overdetermined, linear system as shown in Equations (5.14) and (5.15).

$$\begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3y_3 & y_3^2 & x_3 & y_3 & 1 \\ x_4^2 & x_4y_4 & y_4^2 & x_4 & y_4 & 1 \\ x_5^2 & x_5y_5 & y_5^2 & x_5 & y_5 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & x_ny_n & y_n^2 & x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{5.14}$$

$$AX = 0 \tag{5.15}$$

The (x_i,y_i) data forming the coefficient A matrix are the locations of the corrected edge data on the camera's ideal image plane. The six variables that form the X vector are the unknown ellipse parameters of Equation (5.2). Equation (5.14) is formulated and solved as a polynomial regression problem (Ott, 1993 and Press et al. 1992). The desired result is a set of ellipse parameters that best describe the hundreds of edge data points for a tube-end, in a least-squares sense. The quality of fit in the regression can be assessed both manually and automatically. In the manual approach, the locus of points satisfying the ellipse equation is iteratively mapped back through Equations (4.5) through (4.9) and overlayed on the tubesheet image in the vision board frame buffer. The operator can then compare the regressed ellipse to the original tube-end to ensure the fit is reasonable. The tube-end picture on the far right of Figure 5.3 shows the regressed ellipse overlayed on the tube-end. To automatically quantify the quality of fit, the residuals from the regression process are examined. In addition, some of the edge data can be withheld from the regression and used to investigate the ability of the ellipse equation to predict their locations. This provides a better assessment because the test data is independent of the regression.

5.1.3 Point Correspondence

Now that the edge data has been extracted and regressed to the best-fit ellipses, a correspondence between specific object and image features must be established. The centers of the tube-ends are the most precise and reliable geometric features in the object space. The center locations are known to better than 0.127 mm (0.005 inches) over the entire tubesheet. Considering the manufacturing processes, the center-to-center spacing of adjacent tubes is expected to be much better. In contrast, the diameter of the tube-ends vary, depending on how the tubes were roll-expanded and welded during their initial installation. The shape of some tube-ends may also vary because they were damaged during prior steam generator operations. In addition, blooming characteristics of CCD sensors may dilate highly illuminated and reflective regions such as tube-end edges. The dilation effectively decreases the inside diameters of the tube-ends in the image.

While the tube-end centers are the most accurately located object features, their corresponding image points are not readily available. An enlargement of the circle and circumscribed square image in Figure 4.1 is shown in Figure 5.6. The figure was created with a CAD system and represents the results of an ideal imaging system. The center of the ellipse is defined by the intersection of the dotted major and minor axes. The projected center corresponding to the center of the object circle is given by the intersection of the diagonals of the circumscribed square. As the figure shows, the center of the ellipse does not projectively correspond to the center of the circle in object space. In fact, depending on the external camera parameter values, the projected center may lie anywhere within the boundary of the ellipse and will only coincide with the ellipse center when the object and image planes are parallel.

The projected centers of the object tube-ends must be constructed from the image data. Although the ellipse and object centers do not correspond, tangency is a projective invariant (Rothwell, et al. 1992). In other words, points of tangency between object features correspond exactly to the points of tangency between the imaged features.

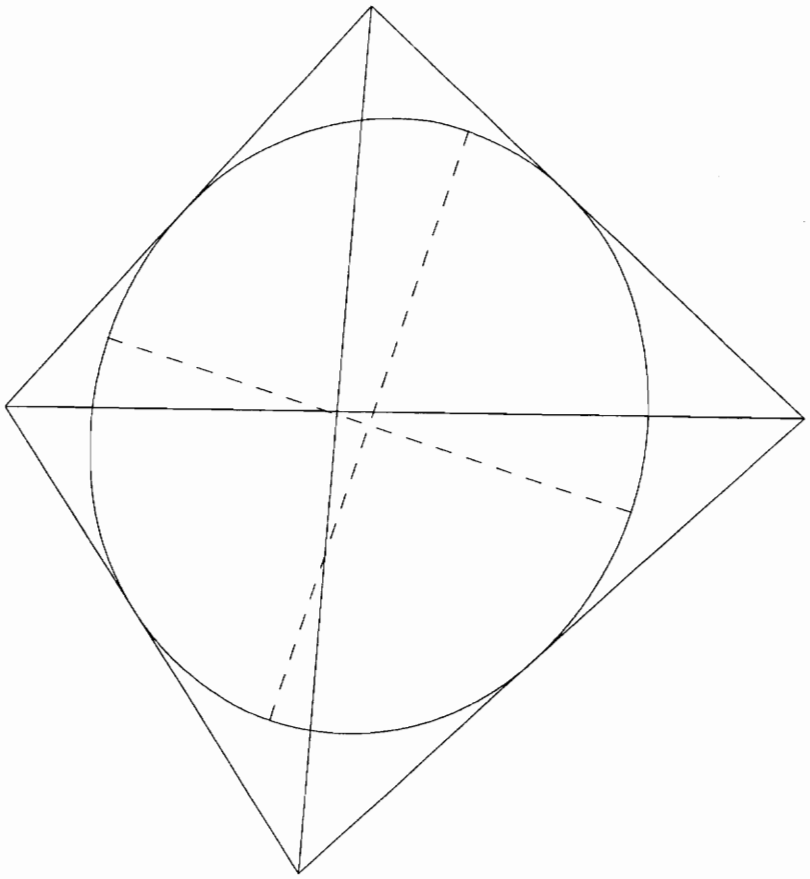


Figure 5.6: Ellipse and Projected Centers on the Image Plane

The preservation of tangency can be shown by referring to Figure 5.7, where line L is tangent to the conic C at point P. The condition of tangency is expressed by the weighted inner product of equation (5.16), where both L and P are expressed as three-tuples.

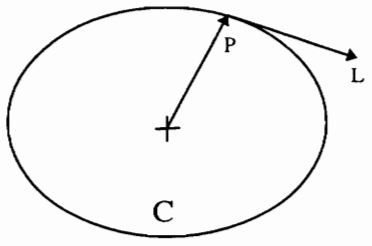


Figure 5.7: Projection Preserves the Tangency of Line *L* to Conic *C*

$$L^tCP = 0 \tag{5.16}$$

In the case of a circle and an ordinary inner product, it is easy to see that the condition of $P \perp L$ implies tangency between the circle and L. In a similar manner, the weighting matrix C enables this condition to be applied to an ellipse. As before, applying a projective transformation T to Equation (5.16) yields Equation (5.17).

$$\left[T^{-1}L'\right]^t C\left[T^{-1}P'\right] = ? \tag{5.17}$$

Using the matrix property of Equation (5.9), Equation (5.17) can be rewritten as Equation (5.18).

$$L'^t\left[T^{-1}\right]^t C\left[T^{-1}\right]P' = ? \tag{5.18}$$

Equation (5.19) is then obtained by substituting a relationship similar to Equation (5.7).

$$[TL]^t\left[T^{-1}\right]^t C\left[T^{-1}\right]TP = ? \tag{5.19}$$

Again using the matrix property of Equation (5.9), Equation (5.20) can be written and then simplified to Equation (5.21).

$$L^tT^t\left[T^{-1}\right]^t C\left[T^{-1}\right]TP = L^tCP \tag{5.20}$$

$$L^tCP = 0 \tag{5.21}$$

By definition, Equation (5.21) is equal to zero, which again implies tangency. Thus, a line and conic that are tangent in object space, are also tangent in their image. Since tangency is preserved, it is easily shown by referring to Figure 5.8 that the tangent points in the image correspond exactly to those of the object. If two conics and their tangent lines are imaged then the points where the lines intersect the conics correspond because, being intersections, both the point on the conic and the line are transformed in exactly the

same manner by T. In the case where just the conics are imaged, then the corresponding lines and points can be constructed in the object and image spaces because tangency is known to be preserved.

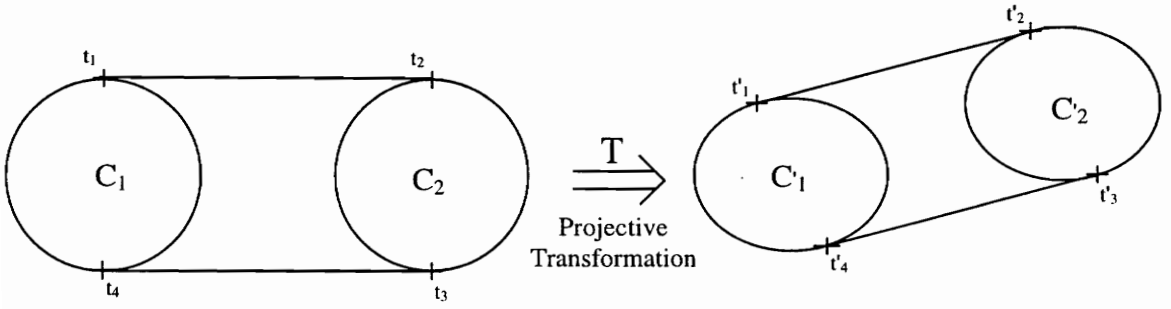


Figure 5.8: Unique Correspondence of Tangent Points

The tangent points between conics can be found in a number of ways including the following three:

1. Intersecting the inverse conic matrices (C_1^{-1} and C_2^{-1}) as described by Rothwell, et al. (1992). This amounts to solving a quartic equation, which yields all four possible tangent lines shown in Figure 5.9.

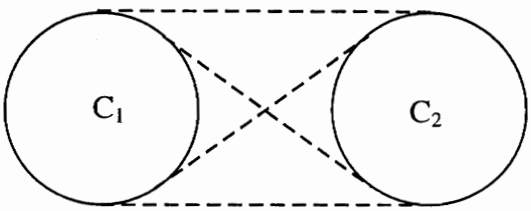


Figure 5.9: Four Bi-Tangent Lines Between Two Conics

As shown in Figure 5.10, a conic can be described either as a locus of points or as an envelope of tangent lines. The point conic is defined by C, while the line conic is defined by C^{-1}. Intersecting the inverse matrices of two conics amounts to finding the four tangent lines that are common to the description of both conics.

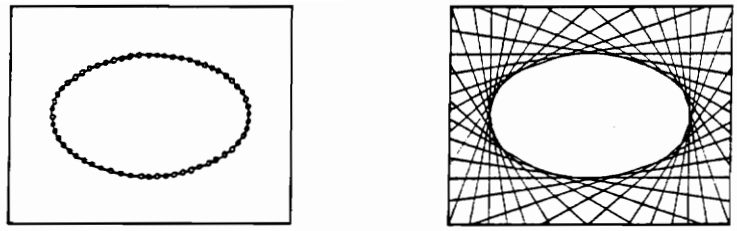


Figure 5.10: Point and Line Conics

2. The second approach consists of using the basic equations enforcing common tangency with a general-purpose iterative solver. This amounts to four unknowns in a nonlinear system of four equations. Two equations constrain common points on the line and two conics, while the other equations constrain the slopes to be the same at those points.

3. Finally, the tangent line between C_2 and an arbitrary starting point on C_1 can be solving for analytically. Then the tangent to C_1 can be solved for, using the tangent point on C_2 as the starting point. This process is then repeated until the common tangent is found between the two conics as shown in Figure 5.11. Since the minor and major axes points are convenient starting points, the iteration normally requires less than five passes to obtain results to the precision of the computer. This approach is used in the work described here, because it is simple and the starting point ensures that only the one tangent line sought is found by the algorithm.

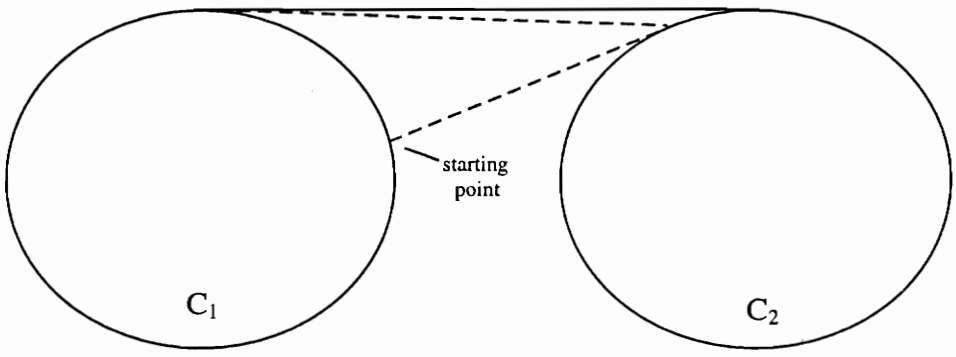


Figure 5.11: Walking Tangent Lines Out to Find the Bi-Tangent Line

Tangent points can now be used to construct the projected centers in the image as shown in Figure 5.12. The centers are recovered by intersecting the joins of opposite tangent pairs. Since the tangent points are preserved under projection, so are the intersections of the joins, by an argument similar to that presented above for tangent points.

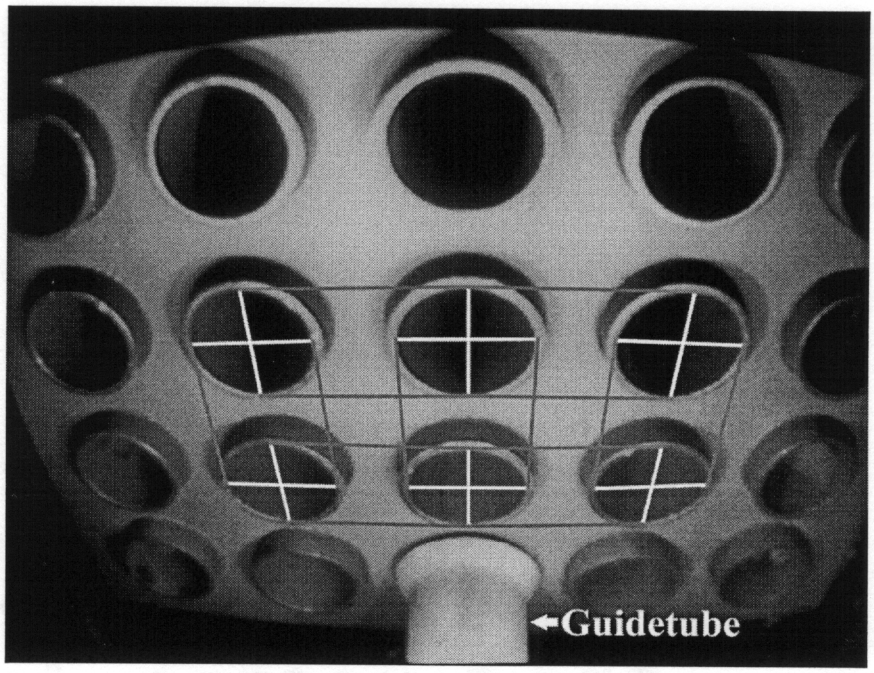


Figure 5.12: Using Tangent Points to Construct Projected Centers

Another interesting method is presented in Figure 5.13, where the projected centers are constructed directly from the tangent lines of neighboring ellipses. Coe et al. (1995) further describes the process of constructing projected centers in the context of the axioms of incidence from projective geometry.

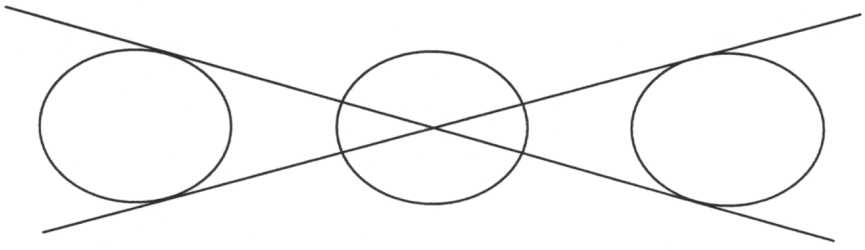


Figure 5.13: Constructing Projected Centers Directly from Tangent Lines

In general, center points are important geometric features to recover in images. Many robotic systems must be able to insert parts or tools into circular holes. Since tool insertion is the primary task, the camera is typically located off-axis and point correspondence must be addressed.

5.2 Estimating Pose Parameters

Now that correspondence has been established between specific point data in the image and object space, the pose between the camera and viewed object can be determined. A variety of solutions exist which generally involve the perspective projection of three, four, and n points. The point data can be the projected centers of tube-ends or other features which may be encoded on guidetube-mounted targets.

5.2.1 Perspective Projection of Three Points

The perspective projection of three points has been studied extensively in photogrammetry (Merritt, 1949) and computer vision (Wolfe et al., 1988 and 1991). The geometry for imaging three object points (p_0, p_1, and p_2) is illustrated in Figure 5.14.

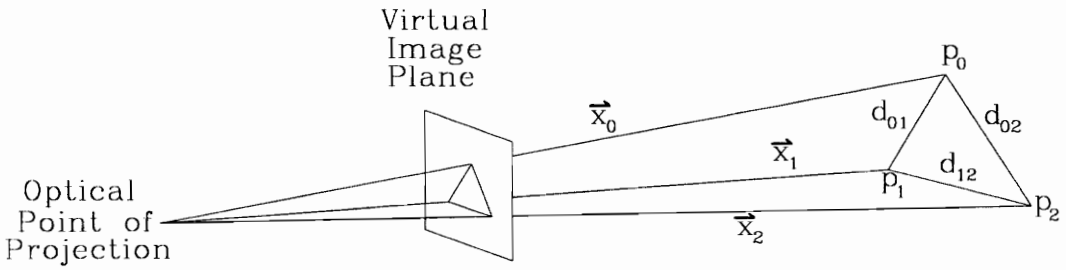


Figure 5.14: Perspective Projection of Three Points

These points form a triangular feature with known relative separations d_{01}, d_{02}, d_{12}. Three rays pierce the camera's image plane as they extend from the optical point of projection to the triangular object. The optical point of projection corresponds to the pin hole of an ideal camera and is located within the lens system of a real camera. The real image plane is located on the CCD sensor inside the camera, which would lie to the left of the optical point of projection in Figure 5.14. A virtual image plane, constructed between the optical point of projection and the object, is used because it is mathematically similar to the real image plane and better suited for visualizing the perspective geometry. The locations of the piercing points on the image plane and the calibrated effective focal length of the camera give the directions (expressed as unit vectors $\hat{x}_0$, $\hat{x}_1$, and $\hat{x}_2$) of the unknown vectors $\vec{x}_0$, $\vec{x}_1$, and $\vec{x}_2$. Using the Law of Cosines three times, the following three nonlinear equations can be written:

$$d_{01}^2 = x_0^2 + x_1^2 - 2x_0x_1(\hat{x}_0 \cdot \hat{x}_1) \tag{5.22}$$

$$d_{02}^2 = x_0^2 + x_2^2 - 2x_0x_2(\hat{x}_0 \cdot \hat{x}_2) \tag{5.23}$$

$$d_{12}^2 = x_1^2 + x_2^2 - 2x_1x_2(\hat{x}_1 \cdot \hat{x}_2) \tag{5.24}$$

The three unknown magnitudes x_0, x_1, and x_2 can be calculated numerically by iteration or the system can be reduced to a fourth degree polynomial and solved in closed-form (Linnainmaa et al., 1988). Although this set of equations has a total of eight solutions, a maximum of four will be positive and real valued. Typically, there will only be two real, positive solutions (Wolfe et al., 1988 and 1991).

An example of two valid solutions for the set of ray magnitudes x_0, x_1, and x_2 are illustrated in Figure 5.15. A valid solution exists wherever the triangle can be positioned and oriented so that its vertices intersect the rays piercing the image plane. In the example shown, this is achieved by rotating the triangle about the axis defined by points p_1 and p_2 until the upper apex intersects ray x_0. The two solutions are given by the

lengths of ray x_0 at the intersection points p_0 and p_0'. Despite the ambiguity, three points can be used if some additional information about the pose of the triangle with respect to the camera is available.

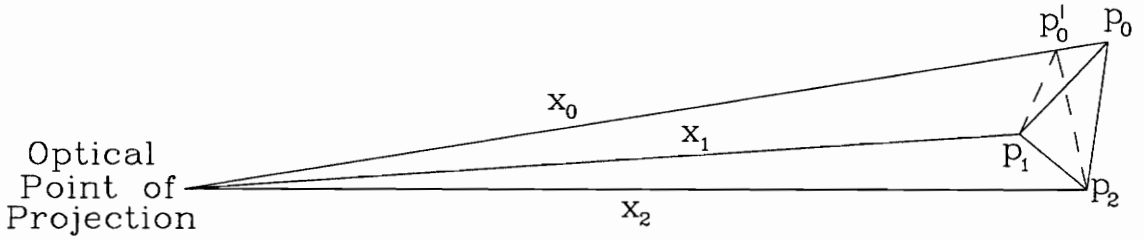


Figure 5.15: Two Valid Solutions for the Perspective Projection of Three Points

Upon solving equations 5.22 through 5.24 and choosing the correct solution, the magnitude and direction of vectors $\bar{x}_0$, $\bar{x}_1$, and $\bar{x}_2$ describe the position of the three object points in space with respect to the camera. The orientation of the triangular feature's reference frame with respect to the camera's reference frame can now be computed.

Figure 5.16 illustrates the camera and an arbitrarily assigned coordinate system (X_t, Y_t, Z_t). Orientation is computed by constructing a rotation matrix that describes the object frame relative to the camera frame. First, the X_t axis is arbitrarily assigned to be along the base of the triangle. A unit vector in this direction is the first column of the rotation matrix as computed with Equation (5.25).

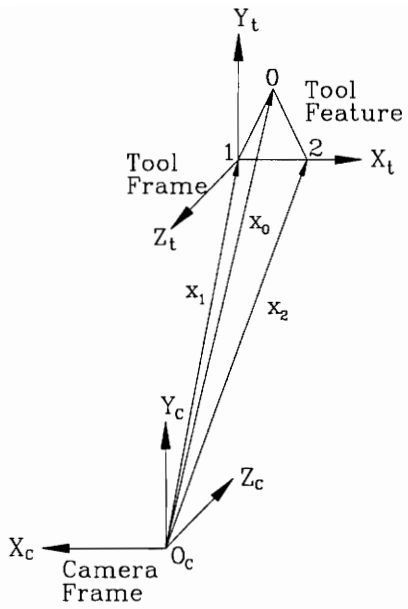


Figure 5.16: Recovering Orientation Parameters

$$\hat{x}_t = \begin{bmatrix} r_{1,1} \\ r_{2,1} \\ r_{3,1} \end{bmatrix} = \frac{\vec{x}_2 - \vec{x}_1}{\left|\vec{x}_2 - \vec{x}_1\right|} \tag{5.25}$$

Second, a vector cross-product between $\vec{x}_t$ and side 0-1 of the triangle is performed and the $\vec{z}_t$ direction defining the third column of the rotation matrix is computed with Equation (5.26):

$$\hat{z}_t = \begin{bmatrix} r_{1,3} \\ r_{2,3} \\ r_{3,3} \end{bmatrix} = \frac{(\vec{x}_2 - \vec{x}_1) \times (\vec{x}_0 - \vec{x}_1)}{\left|(\vec{x}_2 - \vec{x}_1) \times (\vec{x}_0 - \vec{x}_1)\right|} \tag{5.26}$$

Finally, a vector cross-product between the two unit vectors defined by Equations (5.25) and (5.26) is performed and the second column of the rotation matrix is computed with Equation (5.27). This is a unit vector in the $\hat{y}_t$ direction.

$$\hat{y}_t = \begin{bmatrix} r_{1,2} \\ r_{2,2} \\ r_{3,2} \end{bmatrix} = \begin{bmatrix} r_{1,3} \\ r_{2,3} \\ r_{3,3} \end{bmatrix} \times \begin{bmatrix} r_{1,1} \\ r_{2,1} \\ r_{3,1} \end{bmatrix} \tag{5.27}$$

Taken together, these unit vectors define the tool-to-camera rotation matrix:

$$\underset{\text{tool}}{\overset{\text{camera}}{R}}_{x,y,z}(\gamma,\beta,\alpha) = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} \\ r_{2,1} & r_{2,2} & r_{2,3} \\ r_{3,1} & r_{3,2} & r_{3,3} \end{bmatrix} \tag{5.28}$$

Once the rotation matrix is known, a three-angle parameter description of the object coordinate system with respect to the camera coordinate system can be determined with Equations (5.29), (5.30), and (5.31) (Craig, 1989):

$$\alpha = \text{arctan2}\left(\frac{r_{2,1}}{\cos\beta}, \frac{r_{1,1}}{\cos\beta}\right) \tag{5.29}$$

$$\beta = \text{arctan2}\left(-r_{3,1}, \sqrt{r_{1,1}^2 + r_{2,1}^2}\right) \tag{5.30}$$

$$\gamma = \text{arctan2}\left(\frac{r_{3,2}}{\cos\beta}, \frac{r_{3,3}}{\cos\beta}\right) \tag{5.31}$$

where, arctan2 implies the quadrant-specific tangent function

5.2.2 Perspective Projection of Four Points

If more than three points can be used to compute pose, then the problem can be formulated using unconstrained, nonlinear optimization. The ray magnitudes are determined such that they minimize an objective function that is formed from the object's known geometry and its perspective projection on the image plane. Figure 5.17 illustrates the perspective geometry of four points R_0, R_1, R_2 ,and R_3 which form a square on a plane in object space.

The objective function for this example, Equation (5.32), is the sum of the error magnitudes defined in Equations (5.33) through (5.38).

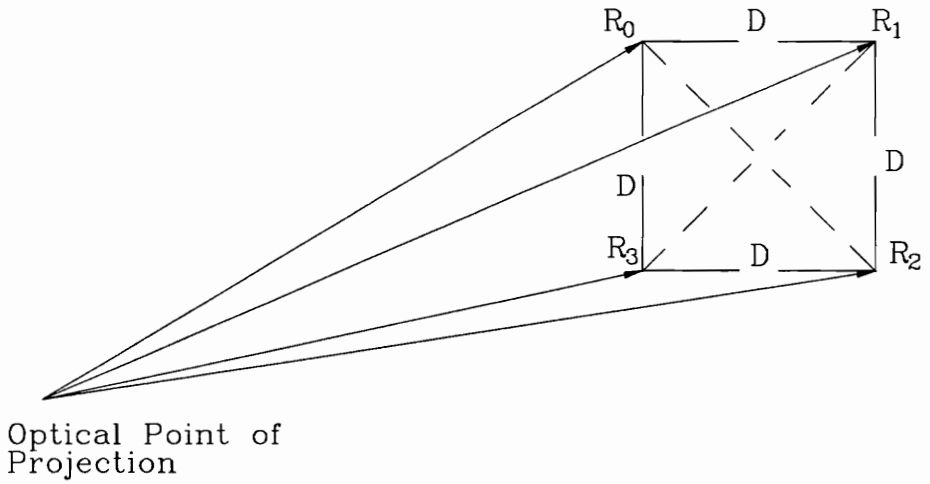


Figure 5.17: Perspective Projection of Four Points

$$\text{Objective Function} = F_0^2 + F_1^2 + F_2^2 + F_3^2 + F_4^2 + F_5^2 \tag{5.32}$$

$$\left|\vec{R}_1 - \vec{R}_0\right| - D = F_0 \tag{5.33}$$

$$\left|\vec{R}_2 - \vec{R}_1\right| - D = F_1 \tag{5.34}$$

$$\left|\vec{R}_3 - \vec{R}_2\right| - D = F_2 \tag{5.35}$$

$$\left|\vec{R}_3 - \vec{R}_0\right| - D = F_3 \tag{5.36}$$

$$\left|\vec{R}_2 - \vec{R}_0\right| - \sqrt{2D^2} = F_4 \tag{5.37}$$

$$\left|\vec{R}_1 - \vec{R}_3\right| - \sqrt{2D^2} = F_5 \tag{5.38}$$

Equations (5.33) through (5.36) establish a quadrangle of known side lengths, while Equations (5.37) and (5.38) constrain the quadrangle to be a square lying in a plane. As with the three point approach, the perspective projection information on the image plane is contained in the directions of the ray vectors. Thus, all of the information known in the object space and measured on the image plane is incorporated into a suitable objective function. The minimization can then be initialized with ray magnitude values computed from multiple three or four point closed-form approaches such as those proposed by Haralick (1989b) and Hung et al. (1985). The solutions from multiple triangle pairs can be matched to remove the ambiguity of three points in perspective (Zeng et al., 1993).

Upon minimizing the objective function, the magnitudes of vectors $\vec{R}_0$, $\vec{R}_1$, $\vec{R}_2$ and $\vec{R}_3$ and the direction information describe the locations of the four object points in space with respect to the camera. The orientation of the square feature's reference frame with respect to the camera frame can then be computed as before.

5.2.3 Perspective Projection of N-Points and Optimization

The four point approaches described and referenced earlier are advantageous because they provide a unique solution, but they are limited to planer points. Three point approaches can be used with more general geometry but give multiple solutions. In both cases, the orientation must be computed separately, which may require a second optimization if the system is over-determined. Another approach is to directly optimize for the spatial transform in which the three position and three orientation parameters are embedded. The ideal relationship between a point in space $[W_x, W_y, W_z]$ and its image $[x_c, y_c]$ is given by Equation (4.1), which is repeated below as Equation (5.39).

$$\begin{bmatrix} h\cdot x_c \\ h\cdot y_c \\ h \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{-1}{EFL} & 1 \end{bmatrix} \cdot \begin{bmatrix} c\alpha\cdot c\beta & c\alpha\cdot s\beta\cdot s\gamma - s\alpha\cdot c\gamma & c\alpha\cdot s\beta\cdot c\gamma + s\alpha s\gamma & x \\ s\alpha\cdot c\beta & s\alpha\cdot s\beta\cdot s\gamma + c\alpha\cdot c\gamma & s\alpha\cdot s\beta\cdot c\gamma - c\alpha\cdot s\gamma & y \\ -s\beta & c\beta\cdot s\gamma & c\beta\cdot c\gamma & z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} W_x \\ W_y \\ W_z \\ 1 \end{bmatrix} \quad (5.39)$$

The process of recovering the spatial transform is described below:

1. The known tubesheet geometry is used to define a local coordinate system and to generate object data corresponding to the image features. The object data is then used as the (W_x , W_y , W_z) points in Equation (5.39).

2. The corresponding image feature points are used as the (x_c, y_c) points in Equation (5.39). Note that the feature points were previously regressed from many image points that were corrected to the ideal pin-hole model described by Equation (5.39).

3. A three or four point method is used to obtain starting values for x, y, z, α, β, γ, and h. The effective focal length may be known from the internal system calibration or it can be included in the minimization process.

4. Using at least four points, an objective function is constructed and minimized to obtain the values of (x, y, z, α, β, γ) which give the best match between the image and object data.

The process outlined above is effective because the starting points for the optimization are the unique results of a reliable closed-form approach. Optimization is then applied to fine-tune the closed-form solution and obtain the best overall transform parameters. Since the minimization begins in the neighborhood of the desired answer, the process is resistant to being trapped in erroneous local minima that may be present because of the system's nonlinearity. Throughout the laboratory and field testing, good convergence has been observed. Optimizing directly for the spatial transform is advantageous because it uses all of the pertinent information known in the object space and measured on the image plane. In addition, the output includes the position and orientation parameters explicitly, it can be applied to general geometries, and is robust in the presence of image noise.

5.2.4 Transform Concatenation

Now that the pose of an object can be determined in the camera coordinate system, two such poses can be combined to compute the pose of the tool relative to the tube of interest. As illustrated in Figures 5.18 and 5.19, the pose of the guidetube tip is first computed relative to the camera frame. This may be accomplished as described in Chapter 4 or by processing the image of a special target mounted at a precisely known location of the tool guidetube.

$^{C}_{GT}T$ = guidetube-to-camera transform

$^{C}_{TS}T$ = tubesheet-to-camera transform

$^{TS}_{GT}T$ = guidetube-to-tubesheet transform

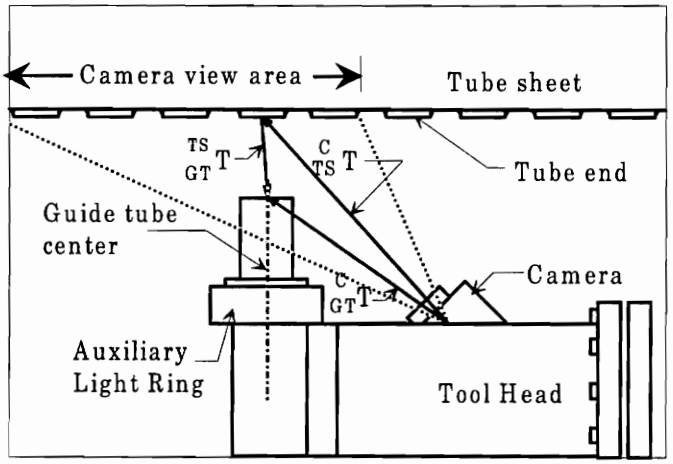


Figure 5.18: Tool-to-Tubesheet Pose

The pose of the tubesheet relative to the camera frame is next computed using the algorithm just described on as may tube-ends in the image as possible (usually six to nine). Finally transformation matrix operations are performed to compute the position and orientation of the tool-tip relative to the center of a particular tube-end, which is the desired result.

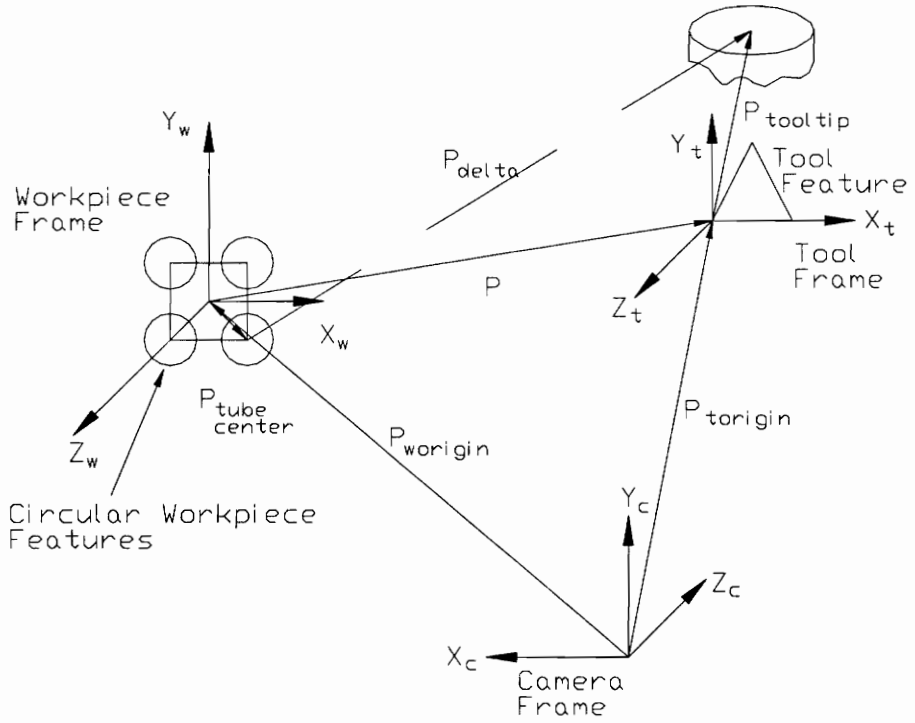


Figure 5.19: Transform Concatenation

Equations (5.40) and (5.41) below give the desired position vector by using the relationships between the tool-tip feature and the actual tool-tip and the relationship between the tubesheet reference point and the center of a tube-end. Equation (5.42) then gives the rotation matrix describing the orientation of the tool frame relative to the tubesheet frame.

$$ {}^{w}P_{tooltip} = \left[{}^{c}_{w}T\right]^{-1} \left\{ \left[{}^{c}_{t}T\right] {}^{t}P_{tooltip} + ({}^{c}P_{torg} - {}^{c}P_{worg}) \right\} \tag{5.40} $$

$$ {}^{w}P_{delta} = {}^{w}P_{tooltip} - {}^{w}P_{tubecenter} \tag{5.41} $$

$$ {}^{TS}_{GT}R = {}^{C}_{TS}R^{T} \; {}^{C}_{GT}R \tag{5.42} $$

Now the complete pose information of the tool-tip relative to the center of the tube-of-interest is available to visually guide the robot.

5.3 Performance Results

The algorithms were simulated in Mathcad (Math Soft, 1995), coded in C, and tested in laboratory experiments. During repeatability tests, the projected center points proved to be stable, with only 60% of the variation observed for the tangent points. In addition using the projected centers reduced the data in a meaningful way for optimization. Measurements with the precision axes system (Figure 3.8) showed positional accuracy on the order of 0.25 mm (0.010 in.). The target accuracy in the horizontal X-Y (tubesheet) plane is ± 0.5 mm (± 0.020 in.) to calibrate the robot base position and ± 1.25 mm (± 0.050 in.) for tool positioning. Target performance in the vertical Z direction is only 5.0 mm (± 0.200 in.), since it is not critical for most steam generator servicing. Robot base calibration, pose verification, and correction were then tested in a full-size steam generator mock-up with the robot control system. The vision system results were compared to traditional robotics approaches under various operating scenarios (Glass et al., 1994b and Fallon et al., 1994a).

Currently, the robot base calibration is performed by docking the robot toolhead over the ends of at least three tubes in different parts of the generator, and using the resulting joint position feedback information to compute the base position. The docking operation must be performed manually by tele-operation. After a good base calibration has been obtained, a large portion of the robot's activities can then be controlled automatically by the computer. Typically, positioning errors are less than 5 mm over the entire generator. More accurate positioning is achieved by adding local zone correction points. These manual docking operations, however, are time consuming, require considerable operator skill, and are accurate to only about 0.5 mm (0.020 in.). In addition, touching the tubes on the tubesheet applies flexural loads to the arm, which are not accounted for in the calibration algorithm. A typical touch calibration for the Cobra robot requires 5 to 15 minutes depending on the operator's skill.

Robot base calibration is performed in a similar manner with the computer vision system, however, the guidetube does not have to be precisely aligned with a tube-end or touch the tubesheet. The operator simply clicks the cursor in the darkened area of four or more tube ends. The task of tagging ellipses requires only a few seconds to complete, in contrast with several minutes for the traditional "touch" approach. A vision-based calibration should be possible within three minutes with almost any operator, since the task is no longer skill dependent.

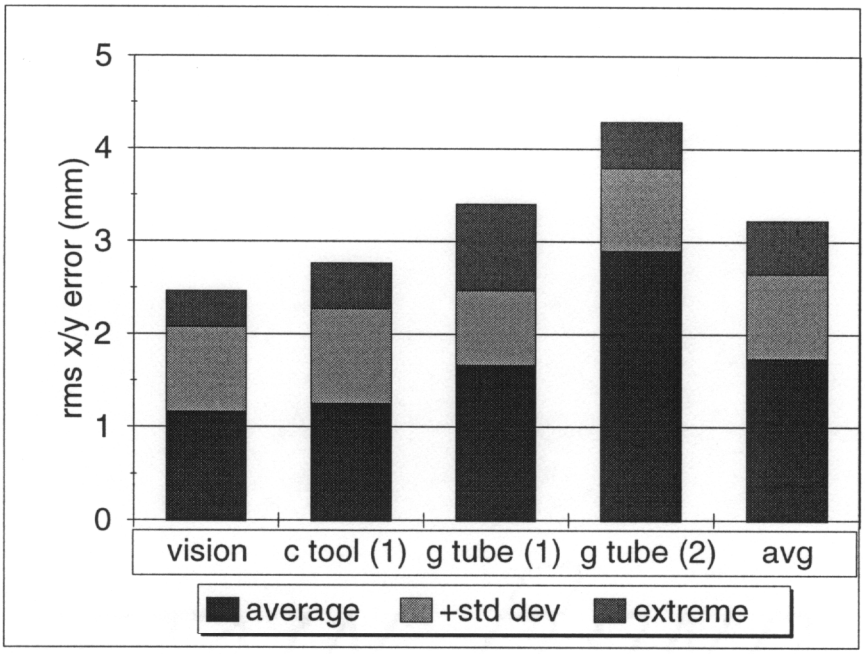
The mock-up testing consisted of comparing four calibration scenarios:

1. the vision system,
2. touch calibration with a highly skilled operator using a special calibration tool,
3. touch calibration with a highly skilled operator using a standard eddy-current tool guidetube, and
4. touch calibration with an average skilled operator using the standard tool.

After each calibration, robot positioning errors were checked at seven tubes distributed over the entire generator using the special calibration tool. The calibration tool fits closely inside the tube diameter such that positioning error was less than 0.25 mm. The calibration tool is designed for testing in the mock-up, but is not used in the field during actual steam generator servicing. During servicing, a standard robot tool with a guidetube must be used to calibrate the robot base. Results for the X-Y radial error are graphed in Figure 5.20, while the vertical Z errors are graphed in Figure 5.21.

As Figure 5.20 shows, the vision system performed better than a highly skilled operator using the special calibration tool and markedly better than a typical operator using a standard tool. This later comparison best represents expected field performance improvements with the vision system. In Figure 5.21, the vision system and highly

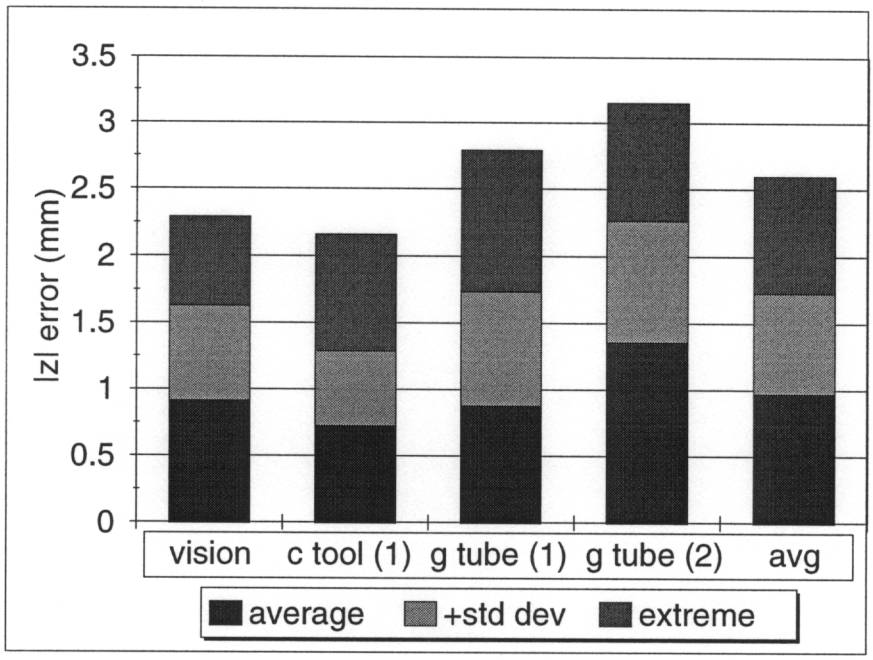
skilled operator are observed to be about the same. Since the Z error is not as critical as the radial error, both results are reasonable.



Notes: (1) touch calibration performed by a highly skilled operator
(2) touch calibration performed by an average skilled operator

Figure 5.20: Radial Error Comparison of Calibration Methods

Calibration for most manipulators includes some zone or local calibration correction feature. As the manipulator moves further away from the initial calibration points, error tends to increase. Local calibration points may be added to adjust the position calculation. During the course of an outage, there may be hundreds of these local calibration points. Each local point may require one to two minutes to define using the touch point approach. Use of the vision system should minimize the required number of calibration points because the base calibration would be better at the outset. Any local points to be added would require only seconds rather than minutes and could be obtained without interrupting the servicing activities.



Notes: (1) touch calibration performed by a highly skilled operator
(2) touch calibration performed by an average skilled operator

Figure 5.21: "Z" Error Comparison of Calibration Methods

Eddy-current probes are rather forgiving in their positioning alignment requirements. Typically, a probe will successfully enter the tube when the alignment error is less than 2-3 mm. Other tools like roll plug, stabilizer insertion, and welded plug tools, however, must be precisely aligned with the tubes to better than 0.5 mm. A local application of the vision system to verify the tool pose and provide a correction when necessary would facilitate precise alignment of these tools as well. Position correction simulations were found to be better than 0.25 mm and enabled smooth tool insertion into the tubes.

The current implementation of the vision calibration system requires some operator interaction to identify tubes. This task can be further automated so that the computer identifies the tubes, calculates the tool alignment error, and implements a corrective move of the manipulator arm. Demands of the robot joints to accurately move a small distance and align with a tube are far less than to accurately locate over a three meter diameter

workspace. Thus, the vision system could permit less costly robots to be developed while still improving positioning performance.

The vision-based approach offers significant advantages over current calibration methods. Since the current manually controlled docking process is replaced by a fast, automatic approach, it is feasible to use more tubes at different tubesheet locations to improve calibration accuracy. Additionally, zone points can be readily added during servicing operations to continually improve the initial calibration model. Finally, the vision-based approach makes it feasible to incorporate more model parameters into the calibration process. The net result is a calibration method that is faster, easier to use, and more accurate than the current approach.

Chapter 6

6. Tube Tracking

The purpose of the vision-based tracking system is to automatically track and identify tube row/column locations of the robot toolhead. The tracking system processes images as the toolhead rapidly moves under the tubesheet. It is used as a coarse positioner for the pose determination system discussed in Chapter 5 and also as a position verifier for the tube-location of the tool guidetube (Fallon et al., 1994b). The verification is independent of the manipulator kinematics and replaces the manual counting process that is currently performed prior to crucial servicing activities (Glass et al., 1994a).

The work described here builds upon earlier research conducted by Brem and Nandhakumar which was also funded by B&W Nuclear Technologies (1993 and 1994). Their system counted tubes moving continuously across the image by tracking two tubes in the image which were closest to a predetermined image point. They divided the tube tracking process into three major stages:

1. Correcting image data for the shape distortion caused by the camera viewing angle. Hypothesizing-and-verifying the location of two tube-ends using a spiral search technique.

2. Finding a better estimate of the location of the two tube-end centers on the image plane given initial estimates of their locations.

3. Displaying the row and column number of the tube closest to a pre-defined point in the video image.

Although their system has substantial merit, it also has several important limitations. The operator must manually compare the toolhead location to that computed by the robot control system. This is further complicated because the tubes being tracked rarely correspond to the tube that the robot toolhead is servicing. In addition, the set-up process is cumbersome and information displayed to the user is somewhat primitive. The new system builds upon the strong points of this earlier research while making specific contributions in the following areas:

1. Specialized toolhead lighting was developed to obtain high image contrast which simplifies image processing and increases reliability.

2. Interlaced image data fields are processed independently to reduce ambiguity caused by camera motion.

3. Tube-end edges are detected with a highly sensitive gradient operator over a reduced search space.

4. Tube-end centers are located with more rapid convergence.

5. A comprehensive imaging model was developed, which includes a full internal vision system calibration.

6. Full-3D pose determination techniques have been developed to calibrate the camera/toolhead relationship remotely.

7. Position estimation techniques have been used to rapidly compute the tube being serviced by the toolhead, even if it is occluded by the tool.

8. Tubes that are radially closest to the best lighting in the image are tracked preferentially.

9. Tubes in the image that the operator selects with a mouse are identified on the on-screen display.

10. Special tube conditions are accommodated such as installed plugs.

11. The vision system is integrated with the robot control system so that the verification process can occur automatically and involve the operator only when a discrepancy occurs.

12. A simple and robust user interface has been developed for operator initialization and interaction.

13. Color-coded row/column, direction, and acquisition status are graphically overlaid on the video image.

14. Software and hardware portability have been incorporated so that the system can be moved from its developmental personal computer platform to a more powerful system or the workstation used for robot control.

The remainder of this chapter describes the system's initialization, image processing, tube-position estimation, user display and communications, and performance results.

6.1 Initialization

After the internal system calibration, the camera is installed in the toolhead and an external calibration is performed inside the steam generator (as discussed in Chapter 4).

The goal of the external calibration is to compute the pose (position and orientation) of the guidetube axis in the camera coordinate system. This information will then be used to compute which tube is above the guidetube during the tracking process.

Upon starting, the vision system software reads the internal and external calibration data files. The operator initializes the system for tracking by performing the following three tasks:

1. Select three tube-ends in the image with the mouse and specify their corresponding row and column number. This establishes the tubesheet row and column directions and the relationship between them. These can vary with robot position and steam generator type.

2. Position and size a search window on the image using the mouse. The window establishes search boundaries where the system will track tubes. This area is a function of image quality, lighting, information displayed on the image by the robot control system, and the location of the guidetube and lighthead in the bottom portion of the image.

3. Use the mouse to select a point near the center of the image search space where the lighting produces the highest contrast between the tube-end and the tubesheet surface. Tubes radially closest to this optimum image point will be preferentially tracked among all those present in the search space. The default location is set as the middle of the search space.

After completing the simple initializing procedure, the tube tracking process (and the steam generator servicing operations) can begin. Additional features that were provided in the Brem and Nandhakumar system are also available, such as adjusting the vision board gain & offset, setting initial values for dynamic edge and region detection thresholds, and displaying tube-end estimates in the search space. These features are

generally only needed for troubleshooting. Displaying tube estimates is particularly helpful in checking that the system is calibrated and initialized properly. Given the initialized tracking tubes, the system computes and displays estimates for the center locations all the other tubes in the search space window. A quick look at the estimates overlaid on the video monitor indicates that the system is set-up properly, interpreting the tubesheet geometry correctly, and ready for tracking. The gain and offset can be used to adjust image brightness and contrast, however, the tool auxiliary lighting provides a more effective way of producing images with high contrast and relatively low noise.

6.2 Image Processing

Since the camera images are interlaced, the even and odd horizontal field lines are acquired at different times. If motion occurs during image acquisition, then the even and odd field pixels are shifted by the distance moved in half the frame rate of the 30 Hz camera (1/60 of a second). If substantial motion occurs, then image processing is best performed on one field set at a time. In the system described here, image processing is divided into two steps. First, the even field lines are processed while the odd field lines are being acquired. Then, information is overlaid on the computer and video monitors while the even field lines are being acquired. This strategy enables the system to process the image acquired by the moving camera at the full rate of 30 frames per second, with only a slight loss of image resolution.

Data from the even field lines must be completely processed in the 1/60 of a second field time. Given this short amount of time and limited computer power, processing must be highly efficient. As the robot moves, tubes continually enter, move around, and leave the image so processing the even fields has been divided into the following two tasks:

1. Computing the centers of two tube-ends being tracked in the current video frame, given their locations in the previous frame as estimates, as illustrated in Figure

6.1 (the camera acquires 30 frames/second). Once detected, tube-end edges are bisected to compute the centers of the two tubes being tracked.

2. If possible, replacing the tracked tube-end that is farthest from the optimum tracking point with a new tube-end that is closest to the optimum point in the image. The optimum tracking point location depends on camera angle, tool location, and lighting. The replacement process consists of mathematically estimating where the closest tubes are located, checking if a tube is actually present near this location, and finding the center of the closer tube. Tubes may not be present in some parts of the tubesheet because of structural supports or because the tubes are plugged and welded flush with the tubesheet.

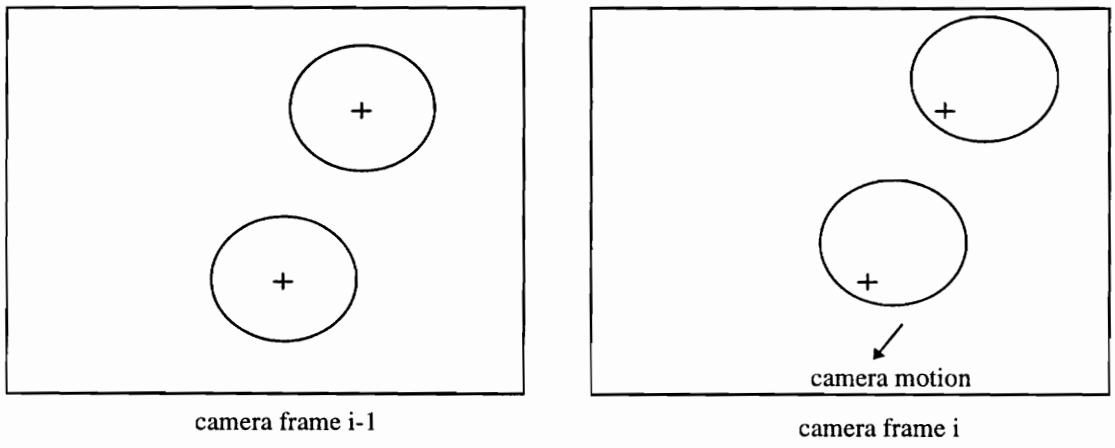


Figure 6.1: Motion in Sequential Images

Estimating tube locations is accomplished with a mathematical imaging model. Tube locations are verified by comparing the intensity sum of a 16 pixel cross-hair to a dynamically adjusted threshold. Tube-ends will have a low sum, while the tubesheet will have a large sum. If a tube-end is not found, a small local search is performed around the prediction. If the local search fails, then the next closest tube-end to the optimum tracking point is examined. Once a closer tube-end is found, then its center is computed from the prediction as described above. As a precaution, the ratio of the tube-end width

and height is inspected to ensure the system did not land in a shadow region near the outside of the tube. To conserve processing time, only one of the two tubes is replaced during a single frame.

After the even-field image processing is complete, the odd fields are processed while the even fields are being acquired from the camera. In addition to displaying data on the computer and RGB video monitors, the odd field processing time is also used to compute which tube is above the guidetube and communicate with the robot control system. Since the tubes being tracked rarely correspond to the tube being serviced, once the tool comes to rest under a tube, the two tubes being tracked are used to estimate which tube is being accessed by the toolhead.

6.2.1 Finding Edges and Centers

Finding tube-end edges and centers is accomplished by considering specified toolhead speeds, tube-end diameters, and camera frame rates. Using this knowledge, it can be shown that the center location estimates from the previous frame will lie between the dark tube-ends of the current frame. Thus, the strategy to locate new tube-end centers is to search for the tube-end edges outward from the center estimates. This is similar to the method discussed in 5.1.1 and repeated in Figure 6.2. First, a vertical search is conducted in the upward and downward image directions from the initial center estimate. After locating the bottom and top edges of the tube-end, the midpoint is used as the initial point for a horizontal search. Then the midpoint between the left and right tube-end edges gives an approximation to the tube-end center in the current frame. How close the approximation is to the actual center depends on the aspect ratio of the ellipse and its orientation. For the steep camera angles in steam generator toolheads, the algorithm's first pass provides a sufficient center estimate for tracking. Additionally, the system continuously cycles through the algorithm to improve the center estimate. Excellent stability and convergence have been observed during both simulation and testing.

Each search consists of two parallel paths which are one pixel wide and separated by one pixel. Edges are found by applying a gradient edge operator to each pixel on the two search paths. The gradient results for corresponding pixels on the two paths are then added together to provide a measure of the edge strength.

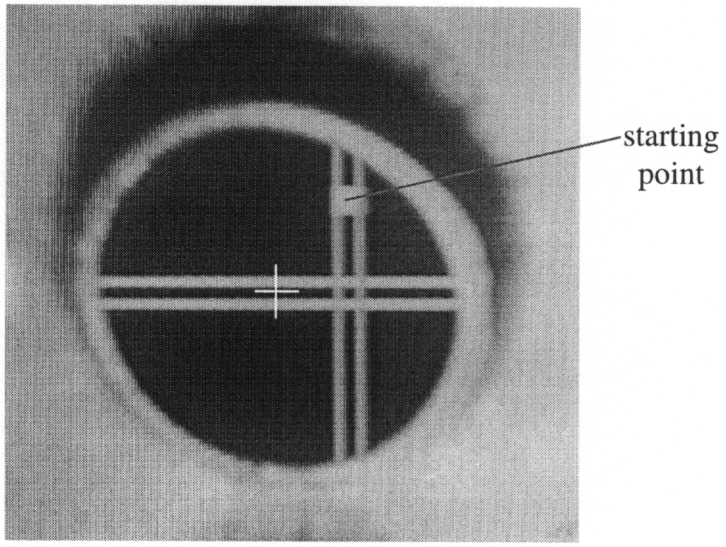


Figure 6.2: Finding the Center of a Tube-End

The horizontal gradient operator uses the intensity values of 24 neighboring pixels on the two paths:

$$\begin{aligned} G_{(u,v)} = {} & I_{(u+1,v+2)} + I_{(u+2,v+2)} + I_{(u+3,v+2)} + I_{(u+4,v+2)} \\ & + I_{(u+5,v+2)} + I_{(u+6,v+2)} - I_{(u-1,v+2)} - I_{(u-2,v+2)} \\ & - I_{(u-3,v+2)} - I_{(u-4,v+2)} - I_{(u-5,v+2)} - I_{(u-6,v+2)} \\ & + I_{(u+1,v-2)} + I_{(u+2,v-2)} + I_{(u+3,v-2)} + I_{(u+4,v-2)} \\ & + I_{(u+5,v-2)} + I_{(u+6,v-2)} - I_{(u-1,v-2)} - I_{(u-2,v-2)} \\ & - I_{(u-3,v-2)} - I_{(u-4,v-2)} - I_{(u-5,v-2)} - I_{(u-6,v-2)} \end{aligned} \tag{6.1}$$

The vertical gradient operator uses the intensity values of 12 neighboring pixels, because only the even fields are processed:

$$\begin{aligned} G_{(u,v)} = {} & I_{(u+2,v+2)} + I_{(u+2,v+4)} + I_{(u+2,v+6)} - I_{(u+2,v-2)} \\ & - I_{(u+2,v-4)} - I_{(u+2,v-6)} + I_{(u-2,v+2)} + I_{(u-2,v+4)} \\ & + I_{(u-2,v+6)} - I_{(u-2,v-2)} - I_{(u-2,v-4)} - I_{(u-2,v-6)} \end{aligned} \tag{6.2}$$

I(u,v) in Equations (6.1) and (6.2) represents the pixel intensity value at column u and row v in the image array, while G(u,v) is the computed gradient value.

The edge and center finding approaches are similar to those described in 5.1.1, except less image data is used for tracking because of the processing time constraints. The approaches facilitate rapid processing by holding the search to a minimum, since it terminates when the gradient value exceeds a dynamically established threshold. The gradient operator produces a large value when an edge is encountered, but is relatively insensitive to image noise, since it depends on many neighboring pixels on two separated paths (Figure 5.2). The center locations are computed rapidly by sequentially bisecting the edge data from the horizontal and vertical searches. Finally, if an edge that is difficult to detect is encountered, the initial search point is shifted in a direction closer to the tube-end center, the dynamic thresholds are reduced, and the search is conducted once again.

6.2.2 Tube Replacement

Given the locations of two tube-end centers on the image plane, a tracking imaging model is used to predict the locations of other tubes in the image for the replacement process described above. Because the tube-ends are axisymmetric, two centers must be continually tracked to extract the necessary position and orientation information.

The model must accommodate a variety of camera mounting positions and orientations while accurately estimating optical and perspective distortions. The extent of the optical barrel distortion is readily apparent in Figure 5.12, where the top row of tubes curves inward. Figure 5.12 also shows the perspective shape distortion, where the distance between the tube-end centers is larger at the top of the image than at the bottom. Since the distortion present in the image is relatively large, a rigorous model must be derived to achieve adequate predictions.

Brem and Nandhakumar (1993 and 1994) present a derivation to correct image data for foreshortening (shape distortion) based on transformation matrices and a two-vanishing point model. This present work derives a complete and more generalized three-vanishing

point model with analytic geometry and a single camera-centered coordinate system. Referring to Figure 6.3, the general methodology is as follows:

1. Remove barrel distortion by mapping the two known tube-end centers in the frame buffer coordinate system to the camera coordinate system using Equations (4.5) through (4.9). The origin of the camera coordinate system is at the optical projection point (a point in the lens system that corresponds to the pin-hole of an ideal camera). To simplify the model, the CCD image plane which lies inside the camera is mirrored to lie in front of the camera and is called the virtual image plane (VIP).

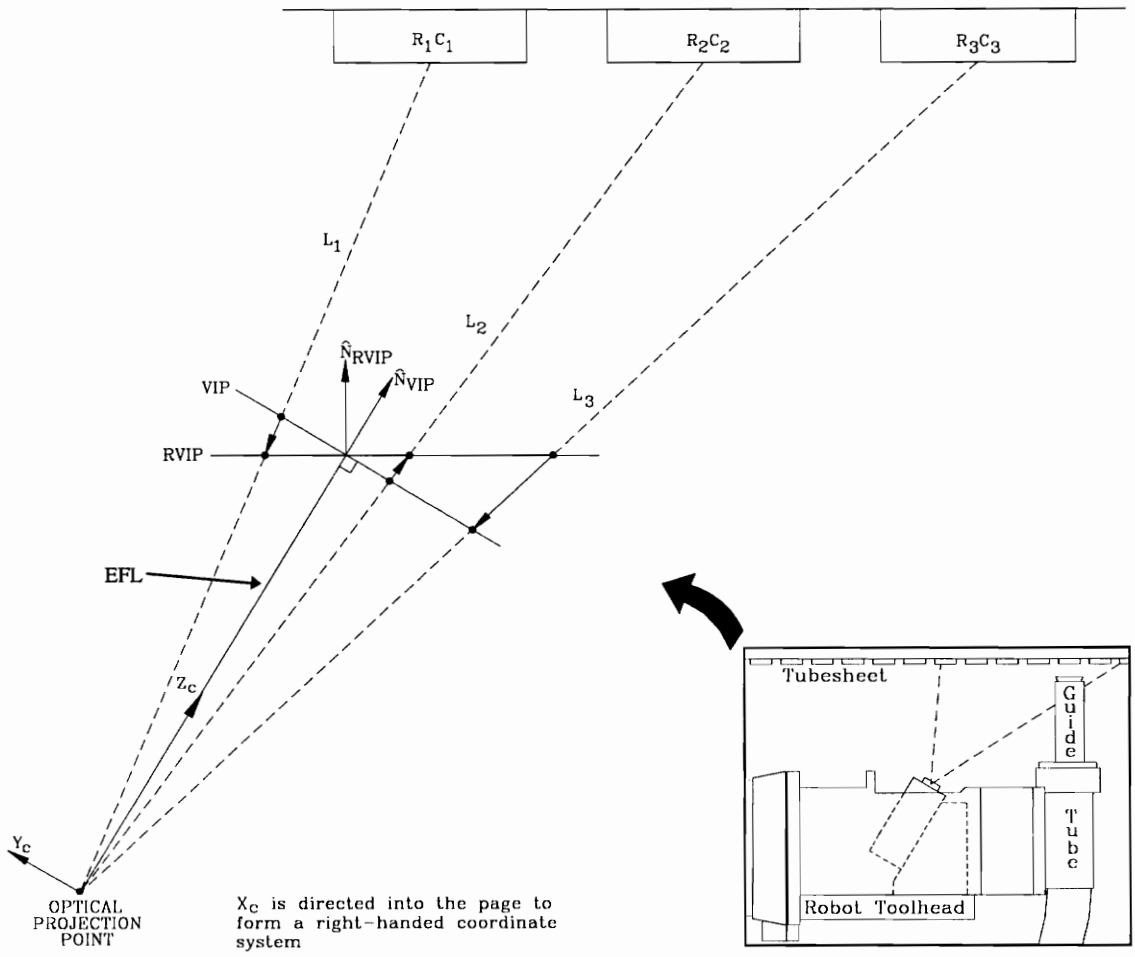


Figure 6.3: Geometry of the Tracking Imaging Model

2. Define ray lines which originate at the origin of the camera coordinate system and extend through the virtual image plane. Compute the intersection points of these lines with a rotated virtual image plane (RVIP) which is parallel to the tubesheet plane. This process removes shape distortion, producing an image that is smaller, but geometrically similar to the actual tubesheet.

3. Use the two known centers of the tubes being tracked and the known geometry of the tubesheet to predict additional tube-end center locations, such as those close to the optimum tracking point and within the search space boundaries. The optimum tracking point and the search space boundaries only have to be defined on the rotated virtual image when they are initialized by the operator.

4. Project the new centers back to the virtual image plane and then map them to the frame buffer coordinate system, where the barrel distortion is restored.

Referring again to Figure 6.3, where a simple case is illustrated, the process is accomplished mathematically as follows:

Define the equation of a plane in space:

$$N_0(x - P_0) + N_1(y - P_1) + N_2(z - P_2) = 0 \tag{6.3}$$

where, the vector $\bar{N} = [N_0, N_1, N_2]^T$ is normal to the plane and

(P_0, P_1, P_2) is an arbitrary point lying in the plane

for the VIP, $\hat{N}_{vip} = [0,0,1]^T$ and $P = (0,0,EFL)$, thus the VIP equation is described as:

$$z = EFL \tag{6.4}$$

for the RVIP, $\hat{N}_{rvip} = [\hat{N}_0, \hat{N}_1, \hat{N}_2]^T$ and $P = (0,0,EFL)$, thus the RVIP description is:

$$\hat{N}_0 x + \hat{N}_1 y + \hat{N}_2 z = \hat{N}_2 (EFL) \tag{6.5}$$

where, $\left[\hat{N}_0, \hat{N}_1, \hat{N}_2\right]$ is determined during the external system calibration process as perpendicular to the tubesheet plane and parallel to the guidetube axis.

Next, lines are described symmetrically as the intersection of two planes in space (Riddle, 1992):

$$\frac{x - P_0}{V_0} = \frac{y - P_1}{V_1} = \frac{z - P_2}{V_2} \tag{6.6}$$

where, the vector $\vec{V} = \left[V_0, V_1, V_2\right]^{\mathrm{T}}$ is between any two points on the line and

(P_0, P_1, P_2) is an arbitrary point lying on the line.

The projection point, which is the origin of the camera coordinate system, will be used as point P and as one of the two points defining vector $\vec{V}$. Thus, known tube-end centers on the VIP can be used to describe lines:

$$\frac{x}{X_{vip}} = \frac{y}{Y_{vip}} = \frac{z}{EFL} \tag{6.7}$$

Solving Equations (6.5) and (6.7) simultaneously gives the point where the ray lines pierce the RVIP in space:

$$Z_{rvip} = \frac{\hat{N}_2 (EFL)^2}{\hat{N}_0 X_{vip} + \hat{N}_1 Y_{vip} + \hat{N}_2 (EFL)} \tag{6.8}$$

$$X_{rvip} = \left(\frac{Z_{rvip}}{EFL}\right) X_{vip} \tag{6.9}$$

$$Y_{rvip} = \left(\frac{Z_{rvip}}{EFL}\right) Y_{vip} \tag{6.10}$$

The projection point and the predicted center location of a tube-end on the RVIP are then used to describe a line:

$$\frac{x}{X_{rvip}} = \frac{y}{Y_{rvip}} = \frac{z}{Z_{rvip}} \tag{6.11}$$

Solving Equations (6.4) and (6.11) simultaneously gives the point where the ray line pierces the VIP in space:

$$Z_{vip} = EFL \tag{6.12}$$

$$Y_{vip} = \frac{EFL}{Z_{rvip}} Y_{rvip} \tag{6.13}$$

$$X_{vip} = \frac{EFL}{Z_{rvip}} X_{rvip} \tag{6.14}$$

Equations (6.8) through (6.10) and (6.12) through (6.14) comprise the tracking imaging model. Although a simple case is illustrated in Figure 6.3 for clarity, the model can accommodate any position and orientation between the image and object planes. Equations (6.8) through (6.10) are used to project two tube-end centers currently being tracked on the VIP to the RVIP, where perspective shape distortion is removed. Since the shape distortion is removed, the known tubesheet geometry can then be used to accurately predict the locations of other tube-ends in the image. After a new tube-end location is found on the RVIP, Equations (6.12) through (6.14) are used to project its location back to the VIP.

Perspective shape distortion occurs because the camera image and tubesheet object planes are not parallel. Figure 6.3 illustrates the two cases of imaging equispaced tube-ends onto non-parallel (VIP) and parallel (RVIP) planes. In the non-parallel case, the distance between lines L_1 and L_2 is greater than the distance between lines L_2 and L_3, as measured on the VIP. The differing lengths indicate that both the image size and shape are distorted on the non-parallel plane. In the parallel case, the distances measured between

the lines on the RVIP are the same, differing only in size from the tubesheet object. Mapping image points to the rotated virtual image plane creates an image that is equivalent to one having no shape distortion because it was made with parallel image and object planes. This concept is further illustrated with a two-stage general proof of the simple case illustrated in Figure 6.4.

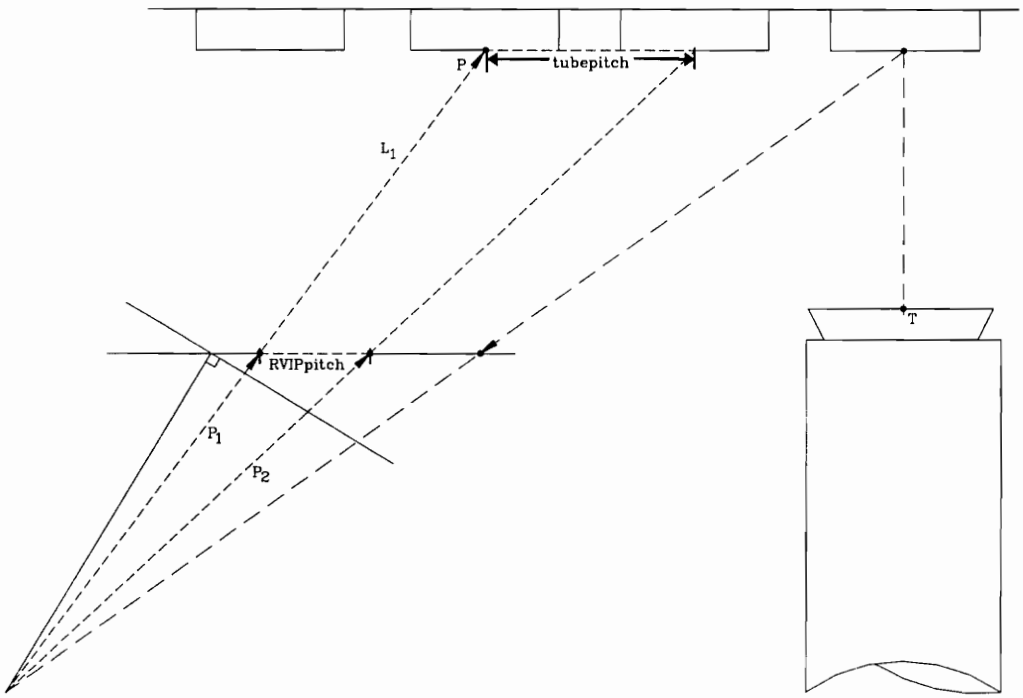


Figure 6.4: Similar Triangle Geometry of the Imaging Model

Since the tube-end spacing is equidistant everywhere in the tubesheet object, it is necessary to prove that the image on the RVIP is also equidistant everywhere. The first step is to use the theorem of similar triangles (Riddle, 1992 and Euclid (Heath, 1912)) to show the following relationship:

$$\frac{\text{tubepitch}}{\text{RVIPpitch}} = \frac{L_1}{P_1} \quad \Rightarrow \quad \text{RVIPpitch} = \text{tubepitch}\left(\frac{P_1}{L_1}\right) \tag{6.15}$$

Since the tubepitch in Figure 6.4 is the same everywhere, the second step is to show that the ratio P_1/L_1 is also the same everywhere. This is accomplished with Thales' Theorem (Berger, 1977) which states that ratio as given in Equation (6.15) between parallel hyperplanes depends only on the spacing of the hyperplanes and not on the location or direction of the line piercing them. Thus, the images of the tube-ends on the RVIP are shown to be equidistant everywhere, even considering tubes very far away from those illustrated in Figure 6.4. Although the proof is illustrated for the simple equispaced case, the generality of the proof extends to any arbitrary object geometry.

6.3 Tube-Position Estimation

As the distance between the guidetube and the tubesheet changes during servicing operations, the tube above the guidetube appears in different positions in the image. Because of this effect of the camera's off-axis viewing angle, robot operators cannot easily verify which tube is actually above the guidetube in a given video image. The vision approach described here enables the tube above the guidetube to be automatically identified independent of the distance that the guidetube is below the tubesheet.

The accuracy and robustness achieved with a complete three-vanishing point formulation further enables the system to estimate position in 3-D space. This is particularly important in steam generator applications where the tube being serviced is often occluded by the tool's guidetube and, therefore, rarely corresponds to one of those being tracked. Once the tool comes to rest under a tube, the two tubes being tracked are used with a 3-D position estimation model to compute which tube is being accessed by the toolhead. Referring to Figure 6.4, the general methodology is as follows:

1. The pitch distance between two tube-ends on the rotated virtual image plane is used with the known tubesheet pitch and the theorem of similar triangles to determine the vector $\bar{P}$ from the projection point to one of the tubes being tracked.

2. The vector $\vec{P}$ is then used with the tubesheet normal vector to define the tubesheet plane in the camera coordinate system. The tubesheet normal vector is known from the external system calibration procedure.

3. Next, the intersection of the guidetube axis and the tubesheet is computed. The location of the guidetube and its axis direction are also known in the camera coordinate system from the external system calibration procedure.

4. Finally, the guidetube axis intersection point is projected to the rotated virtual image plane and compared with undistorted tube-end center locations to determine which tube is actually above the guidetube.

Referring again to Figure 6.4, where a simple case is illustrated, the process is accomplished mathematically as follows:

Define vector $\vec{P}$:

$$\vec{P} = \frac{tubepitch}{\left|(\vec{P}_2 - \vec{P}_1)\right|}\vec{P}_1 \tag{6.16}$$

Where, $\left|(\vec{P}_2 - \vec{P}_1)\right|$ is the pitch on the RVIP and

$\vec{P}_1$ extends from projection point to the RVIP (from the tracking imaging model)

Define the equation of the tubesheet plane in space:

$$\hat{N}_0(x - P_0) + \hat{N}_1(y - P_1) + \hat{N}_2(z - P_2) = 0 \tag{6.17}$$

Describe the guidetube axis with a symmetric line equation in space (Note that since the guidetube axis is perpendicular to the tubesheet, the tubesheet normal is used again):

$$\frac{x - T_0}{\hat{N}_0} = \frac{y - T_1}{\hat{N}_1} = \frac{z - T_2}{\hat{N}_2} \tag{6.18}$$

Where, (T_0, T_1, T_2) is a point defined by the external system calibration procedure as the tip of the guidetube.

Solving Equations (6.17) and (6.18) simultaneously gives the point where the guidetube axis intersects the tubesheet object plane:

$$Z_o = \frac{\frac{\hat{N}_0^{\,2}}{\hat{N}_2} T_2 - \hat{N}_0 T_0 + \hat{N}_0 P_0 + \frac{\hat{N}_1^{\,2}}{\hat{N}_2} T_2 - \hat{N}_1 T_1 + \hat{N}_1 P_1 + \hat{N}_2 P_2}{\frac{\hat{N}_0^{\,2}}{\hat{N}_2} + \frac{\hat{N}_1^{\,2}}{\hat{N}_2} + \hat{N}_2} \tag{6.19}$$

$$X_o = \frac{\hat{N}_0}{\hat{N}_2}(Z_o - T_2) + T_0 \tag{6.20}$$

$$Y_o = \frac{\hat{N}_1}{\hat{N}_2}(Z_o - T_2) + T_1 \tag{6.21}$$

Next, a line is defined from the projection point though the point described by Equations (6.19) through (6.21):

$$\frac{x}{X_o} = \frac{y}{Y_o} = \frac{z}{Z_o} \tag{6.22}$$

Finally, solving Equations (6.5) and (6.22) simultaneously gives the point where the line pierces the RVIP in space:

$$Z_{rvip} = \frac{\hat{N}_2 (EFL) Z_o}{\hat{N}_0 X_o + \hat{N}_1 Y_o + \hat{N}_2 Z_o} \tag{6.23}$$

$$X_{rvip} = \left(\frac{Z_{rvip}}{Z_o} \right) X_o \tag{6.24}$$

$$Y_{rvip} = \left(\frac{Z_{rvip}}{Z_o} \right) Y_o \tag{6.25}$$

The point defined by Equations (6.23) through (6.25) is compared with tube-end center locations on the RVIP to determine which tube is actually above the guidetube. RVIP tube-end center locations are constructed with the tracking imaging model using the two tubes being tracked and known tubesheet geometry. As the distance between the guidetube and the tubesheet changes during servicing operations, the tube above the guidetube appears in different positions in the image. The approach described here enables the tube above the guidetube to be correctly identified independent of the distance that the guidetube is below the tubesheet.

Both the tracking and tube-position models accommodate any tubesheet geometry and camera mounting. Examples of different tubesheet geometries include triangular-pitch, flush-welded pictured in Figure 6.5, square-pitch, flush-welded pictured in Figure 6.6, and square-pitch, fillet-welded pictured in Figure 5.12. In addition, the cameras in Figure 6.6 and Figure 6.5 are mounted with orientations differing by 90 degrees about their optical axes.

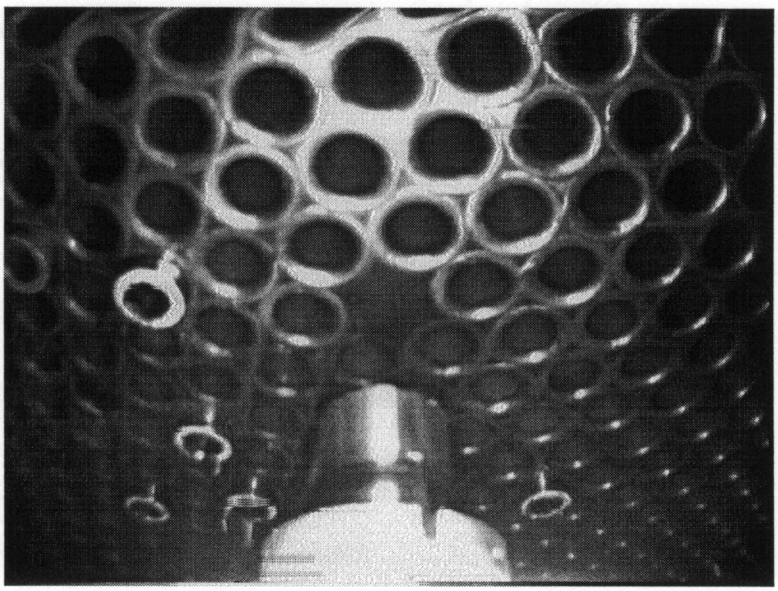


Figure 6.5: Triangular-Pitch, Flush-Welded Tubesheet

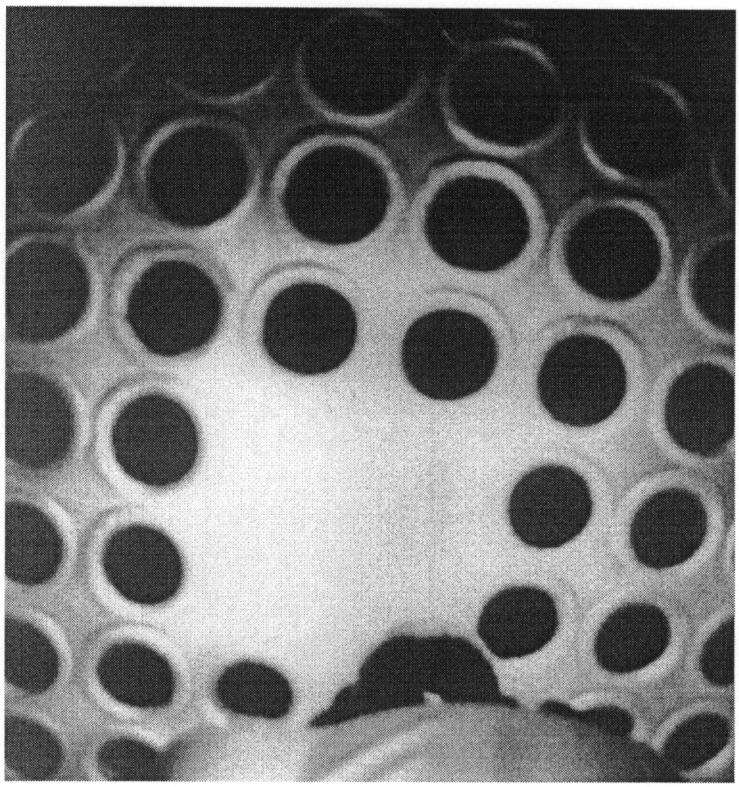


Figure 6.6: Square-Pitch, Flush-Welded Tubesheet

6.4 User Display and Communication

The information displayed on the RGB video monitor during the tracking process is illustrated in Figure 6.7. The two tubes being tracked are near the center of the image. The center of the primary tube-end is indicated by the origin of an L-shaped row/column coordinate system marker laid over the image data. The center of the second tube being tracked is indicated by a small cross-hair. As shown in the photo, the row vector which points along the positive row direction is colored red and the column vector which points along the positive column direction is colored yellow. The row and column number corresponding to the primary tube are also displayed with appropriate colors in a black box at the bottom of the image. Since writing large numbers to the frame memory for display requires a substantial amount of time, row and column numbers are updated during alternate frames.

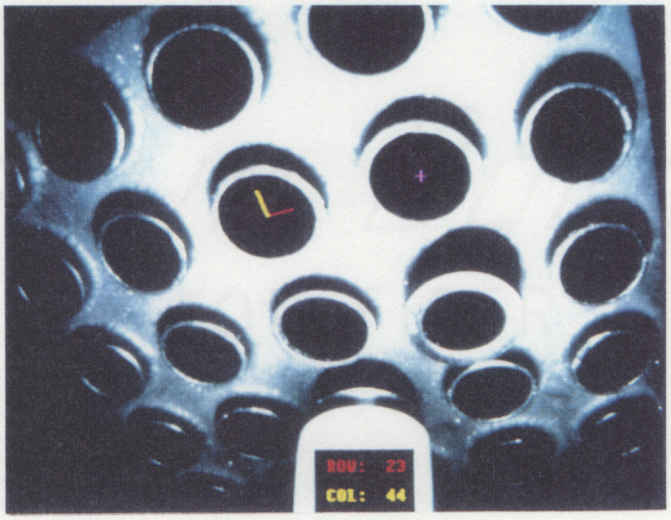


Figure 6.7: Vision System Color Overlays on Video Monitor

Note that the primary tube is chosen as the best tube for the vision system to track and rarely corresponds to the tube above the guidetube. If the vision system is queried by the operator through the keyboard or the robot control system through the serial port, the tube above the guidetube is identified with a large cross-hair and its row and column are displayed in a dark blue color. Figure 6.7 also shows a plug installed in a tube-end just to the right of the image center. The plug protrudes below the other tube-ends and is characterized by a depressed central portion and thick edges. During installation, a tool is inserted into the depressed region and the plug expanded and sealed against the tube wall.

The basic operational flow of the vision-based tracking system is illustrated in Figure 6.8. First camera parameters are loaded, which include effective focal length, lens distortion, and other internal system parameters as described in Chapter 4. The tool-to-camera pose parameters are then calibrated and loaded as described in Chapters 4 and 5. Next the system is initialized by the operator who selects three tubes and a search space with the mouse, and then enters the tube row and column numbers through the keyboard. Once tracking begins, the system displays the current row and column position of the primary tube being tracked. When the robot control system stops to access a tube, it will then query the position of the guidetube for verification.

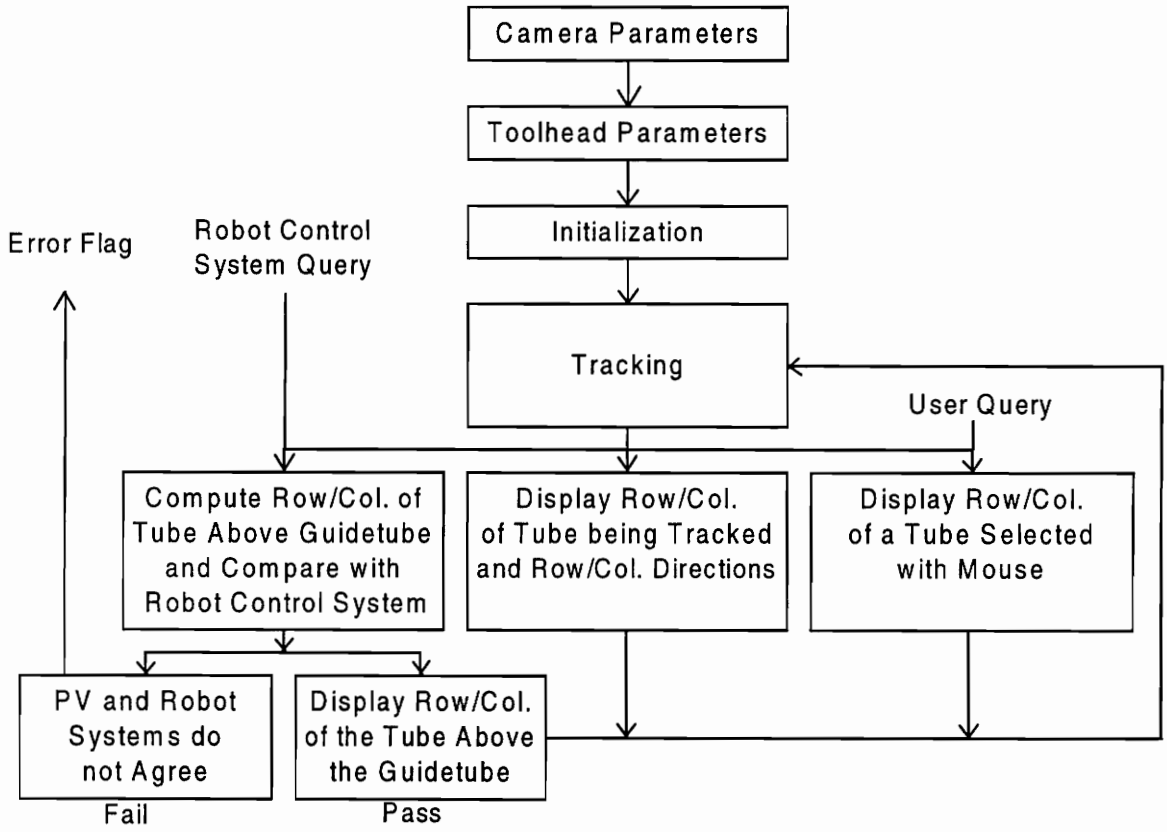


Figure 6.8: Operational Flow of the Tracking System

The system is designed to report the tube above the guidetube only when the guidetube has sufficient alignment to insert an inspection probe. This precaution prevents false discrepancies between the two systems that can be caused when the guidetube is between tubes. In normal operation, the robot control system will send the row and column number of the tube above the guidetube across an RS-232 serial communication port. The vision system will receive the data and compute the row and column independently. If the systems do not agree, the vision system will generate an error flag and the operator can use the graphical information on the RGB monitor to troubleshoot the error. The system operator may also query for either the tube being accessed by the guidetube or any other tube that is selected with a mouse.

6.5 Performance Results

Tracking and position estimation models were described using simple cases only to facilitate their presentation. Both models are general, accommodating any tubesheet geometry and using any two tubes in the camera's field of view, not just adjacent tubes. The models are based on the following assumptions and constraints:

1. the toolhead guidetube remains relatively perpendicular to the tubesheet while the robot services the steam generator,

2. the difference between the center of the ellipse in the tube-end image and the projected center as described in Figure 5.6 is small relative to the tracking requirements,

3. the tube-end pitch spacing is accurately known,

4. all the tube-ends in the tubesheet lie in a plane,

5. once externally calibrated, the camera does not move relative to the guidetube, and

6. once internally calibrated, the camera focus remains fixed.

Computer simulations and sensitivity analyses were performed to investigate the impact of reasonable discrepancies in the assumed values on the models and the performance of the system. The goal of the tracking model is to estimate the locations of tube-ends on the image plane as close to the their center as possible, but no greater than the minimum inside radius away from the actual center. The goal of the position model is to accurately estimate the location of the closest tube-end above the guidetube as accurately as possible and no greater than half a tube pitch away from the actual center. Considering these performance requirements, the results of the analyses lead to the following conclusions:

1. Since the robot must perform insertion tasks, it is carefully designed and mounted to maintain perpendicularly between the guidetube and tubesheet. Small perpendicularly errors which may be present do not substantially impact the system performance.

2. Typically, the difference between the image ellipse center and the projected tube-end center is less than half of a pixel on the CCD imager. While this error cannot be tolerated in the external calibration process, it is easily acceptable for the tracking and position estimation models.

3. The tube-end pitch tolerance is less than 0.127 mm (0.005 inches) over the entire tubesheet and is probably much smaller in the neighborhood of tubes within the camera's field of view. Once again, this small error does not significantly impact system performance.

4. Tube-ends lying in a plane is an acceptable assumption unless a tube-end has been plugged. Generally, the plugs protrude out of the tube-end by varying amounts, depending on the type of plug and how it was installed. While plugs can be tolerated in the tracking model, too much error is accumulated for reliable position estimation. Since it would be very difficult to accommodate the erratic plug installation geometry, they must be detected and avoided by the system when the tube above the guidetube calculation is performed.

5. The camera is rigidly mounted to prevent any motion relative to the guidetube.

6. The camera focus adjustment has a locked-down feature which is used to preserve its calibration.

The computer simulation, sensitivity analysis, and system performance were further verified with a set of experiments and tests with mock-ups and actual steam generators. The laboratory experiments consisted of estimating groups of 11 to 19 tube-end centers at

several camera angles and positions which are typically encountered in steam generator applications. The error is expressed as the following normalized value:

$$\text{Tracking Model Error} = \frac{\text{RMS[(actual center location in frame buffer) - (center estimate)](100)}}{\text{(minor radius of tube-end ellipse in frame buffer)}} \tag{6.26}$$

As long as the error is less than 100%, the model produces a successful estimate because it is within the tube-end boundary. The tubes were assigned numbers according to their position in the image with the first being in the center, and the rest moving outward in a clockwise direction. The results of the three tests are summarized in Table 6.1 and shown in Figure 6.9.

Table 6.1: Test Configurations

	camera angle	tubesheet angle	camera distance	error range
test 1 (foreground)	50 degrees	0 degrees	125 mm	0.5-12.6%
test 2 (middle)	50 degrees	45 degrees	150 mm	0.2-12.8%
test 3 (background)	50 degrees	0 degrees	150 mm	1-23.5%

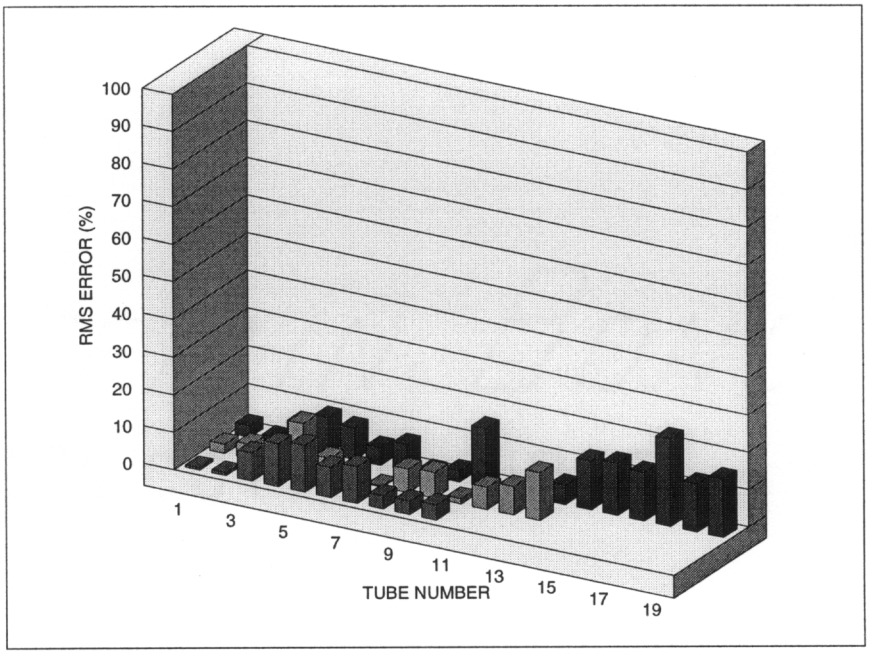


Figure 6.9: Model Accuracy for Typical Cases

Another test was performed which illustrates the general case of three angles of rotation between the camera and the tubesheet. Although this case is not commonly encountered in steam generator applications, it validates the model's capability of handling all three vanishing points. Having this capability provides flexibility in camera mounting and is more accurate than modeling just two angles of rotation. The final experiment produced errors ranging from 1.5 to 15 percent as illustrated in Figure 6.10.

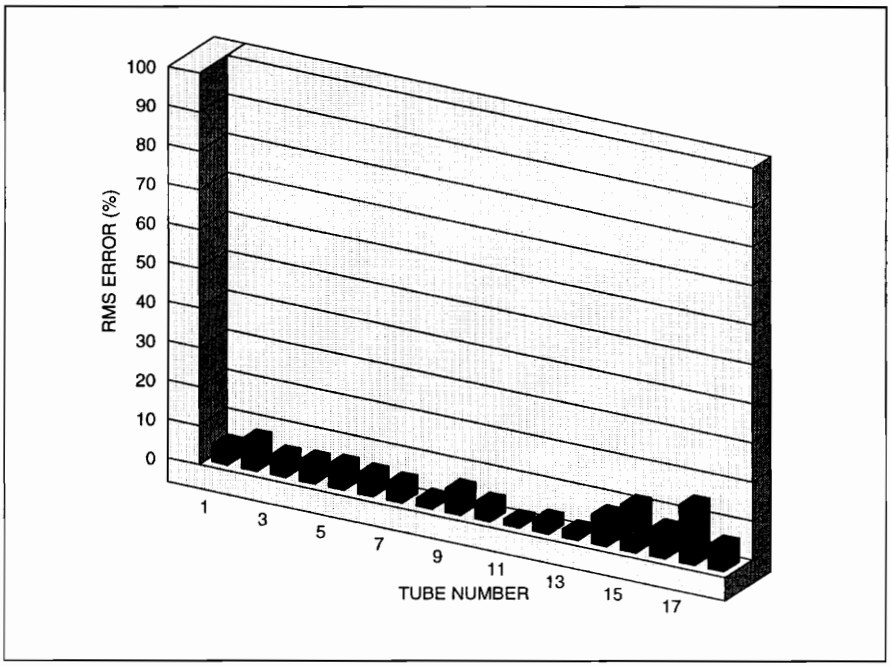


Figure 6.10: Model Accuracy for an Extreme Case

The small errors reported are easily tolerated by the system, since in all cases the models predicted the tube-end centers with a comfortable margin. In most cases the error on the image plane amounts to only a few frame buffer pixels. For some of the smaller tube-ends in the image, however, a few pixels can be 10-20 percent of the tube-end radius. In practice, the system chooses the best tube among all those available in the image based on tube-end location and lighting criteria. If an estimate is outside the actual tube-end (error > 100 percent), the system would simply use a better estimate among those available. Testing the position estimation model also yielded excellent results while computing the

tube above the guidetube in 3-D space. Using the same criteria as above, errors ranged between 1.4 and 14 percent. The error observed for both models is on the order of that predicted by the computer simulations. The largest contributor to the total error comes from computing the ellipse centers on the image plane with a bisection method which uses only a small amount of image information. The bisection method enables the system to processes 60 tubes/sec for toolhead speeds up to 250 mm/sec (10 in/sec) with minimal computer power. Table 6.2 lists the major processing operations and the profile times for the 16.5 msec odd and even field acquisition rate (using a 486-33mhz PC).

Table 6.2: Time Estimates for All System Operations

Even Field Processing Time (16.67ms max)	msec
find the new centers of two tubes	< 8.6
predict tube closer to optimum track point	< 0.5
check if the predicted tube can be processed	< 0.4
find the center of the predicted tube	< 4.3
total	< 13.8
Odd Field Processing Time (16.67ms max)	msec
update display on RGB video monitor	< 13.5
compute tube above guidetube	< 0.5
communicate with robot control system	< 2.0(est.)
total	< 16.0

The vision system has been field tested in mock-ups and actual steam generators at nuclear power plants. During recent tests, the system successfully tracked over 250 consecutive tubes in a typical inspection scenario. The toolhead was moved from tube to tube with a short pause at each tube as if performing about six hours of eddy-current inspection. The system also successfully tracked dozens of long moves (as many as 30 tubes) with the manipulator moving at its highest speed. Heights during these moves varied between 50 mm (2 inches) and 80 mm (3.25 inches) below the tubesheet. For

distances between 0 mm and 90 mm (3.5 inches) below the tubesheet the system could successfully determine which tube was above the guidetube, provided that the guidetube was aligned well enough with the tube to feed the eddy-current probe.

The system is also designed to accommodate many situations encountered in different types of steam generators, such as voids and plugs. Voids are places in the tubesheet where there are no tube penetrations, often because a structural support (stay) is located on the top side of the sheet. The vision system handles voids by tracking tubes around their periphery in the image. Flush-welded plugs do not cause difficulties because they are treated as small stays by the system. Rolled plugs appear as an open tube-end and also provides no tracking difficulties. Rolled plugs which protrude out from the tube-end substantially, however, can cause inaccuracies in the 3-D model. To minimize inaccuracies, plugs need to be detected and avoided when computing 3-D information whenever possible. Rolled plugs can be detected by comparing the wide upper edge thickness to that of a typical tube-end for a fillet sheet (see Figure 6.7 where the upper edge of plug is thick compared to a tube-end). In the case of a flush sheet, plugs are detected by the existence of a finite upper edge thickness which is caused by the plug shadow (see Figure 6.5).

The automated tracking system offers significant advantages over current manual verification methods. Reliability can be improved by automatically comparing the vision system results with the robot's control system. Operator performance can be improved by eliminating the distraction and monotony of manual counting. Finally, the robot can operate faster and with fewer errors. The new tracking system should reduce inspection and repair time, improve quality, and result in safer, less expensive service.

Chapter 7

7. Conclusions and Future Work

7.1 Results

This dissertation described the development of a computer vision system for improving the kinematic control of robots. The vision approach was successfully applied to increase the automation and accuracy of nuclear service robots. The system is divided into its two major functions of pose determination and tracking. Pose determination is used to calibrate the base location of the robot, verify the tool pose for insertions, and compute a precise correction, if necessary. Tracking is used as a coarse positioner and to verify the tool's tube-position prior to performing crucial servicing operations. The vision-based approach is advantageous because it determines position and orientation (pose) parameters more directly than conventional approaches that are based on kinematics and joint feedback.

Mathematical models were developed to accurately describe geometric relationships between objects in space and digital representations of their image. They were shown to predict the location of image features while tracking and also to estimate the position of the robot toolhead once it came to rest at a prescribed servicing location. The vision system was simulated and implemented with the actual robot system used inside steam

generators. The mathematical models prove to work extremely well in computer simulations, carefully controlled laboratory experiments, and actual field tests.

Highlights of the tracking system are summarized as follows:

1. The system processes 60 tubes/sec for toolhead speeds up to 250 mm/sec.
2. Tracking and position estimation models were verified in lab tests.
3. 250 tubes were tracked while simulating six hours of eddy-current inspection inside an actual steam generator.
4. Long moves (up to 30 tubes) were handled at high toolhead speeds.
5. The tube above the tool guidetube was correctly identified, even with varying tool heights.

Highlights of pose determination system are summarized as follows:

1. System performance exceeded the design goals of 0.5 mm for calibration and 1.25 mm for pose determination.
2. Results better than 0.25 mm were produced in laboratory and full-scale mockup tests.
3. Vision-based robot base calibration was successfully demonstrated which was better than a highly skilled operator using a special calibration tool and markedly better than an average operator using an actual servicing tool.
4. Position verification and correction for precision insertion tasks was successfully demonstrated.
5. The same low-cost cameras that are currently used for operator feedback were utilized and they did not have to be precisely mounted.

6. Optimization techniques were implemented, which are particularly well suited because they use all of the pertinent information known in object space and measured on the image plane, can be applied to complex geometry, and are robust in the presence of image noise.

7. The conventional error stack-up in robots was replaced with a much smaller and localized list consisting of vision system calibration parameters; tool and workpiece feature geometry; and pixel quantification and lighting.

The system is also cost effective and compatible with many of the different robots, toolheads, and steam generators found in both domestic and foreign power plants.

While this dissertation presents the background, research objectives, solution strategy and implementation in a highly positive context, many difficulties were faced along the way. Some of these challenges were described in detail as design considerations in Chapter 1 and many others were presented throughout the remaining chapters. The most important challenges were hardware limitations and design constraints. Practical limitations were placed on the size and cost of cameras and on the image processing hardware. The system was further required to function with dozens of existing robots, toolheads and cameras. Additional design constraints arose because the vision system is only one subsystem among several others which interact and form a complex array of technology. Many design trade-offs had to be managed, far more than if success was measured by computer simulations or laboratory investigations alone.

Beyond the present application, this work finds further relevance in many manufacturing and service settings. There is an ever increasing need to evolve technology that has an impact in the global market as well as accomplishing large-scope, specialized tasks such as those arising in environmental remediation. The work described here can contribute to many areas such as construction, maintenance, material handling, welding, spray coating, assembly, inspection, and recycling tasks.

Computer vision is an important technology in the evolution of robotics. It offers powerful sensing capabilities that can be used to improve the adaptability, flexibility, and cost of robot systems. The methodology presented in this dissertation is significant because it provides a more direct measure of the end-effector pose parameters than can be achieved using joint feedback methods. Single-view monocular camera images can be used with geometric knowledge of objects, features, or targets in the workspace to recover the end-effector/workpiece pose relationships. This methodology is practical, because the required geometric knowledge is often available in manufacture and service settings. Although computer vision involves solving difficult theoretical and practical problems, the gains are large, particularly in the field of robotics. Once the proposed methodology is fully implemented and mature, the number of potential applications is enormous.

7.2 Implications

Improving the control of robots with computer can be divided into four levels, depending on how the visual information is utilized. The first level is as an open-loop observer in which the visual information is used to measure the robot tool location without providing any feedback. Examples of this method are calibrating the robot base location and verifying the tube-position of the tool guidetube in the steam generator example.

The second level uses the vision system as a supplement to existing kinematic control. Pose determination and correction is an example of this control strategy in the steam generator example. The corrections can be implemented as new destinations for the end-effector in the workspace.

The third level is an incremental-kinematic approach that eliminates the need for joint feedback. In this approach the robot kinematic model is retained and used to move the robot based on the vision system measurements. Replacing joint feedback is advantageous

because many wires (28 for the Cobra steam generator robot, for example) and costly high resolution joint encoders could be eliminated.

Finally, the fourth level of control strategy is to replace the kinematic model with an intelligent algorithm that does not depend on knowledge of the robot's configuration or exact geometry. This approach can begin with only basic knowledge about the robot system which improves over time with a learning process. Although intelligence-base control is a formidable problem, it offers advantages in both the flexibility and cost of robots. Some preliminary study with a general-purpose PUMA 560 robot indicates that this is possible, even for complex robot geometries (Robotics and Mechanisms Group Report, Virginia Tech, 1994).

Current robot control systems are specialized because they depend on known robot and tool geometries as well as substantial knowledge about the workspace and workpiece. Often the controller requires not only that the robot geometry be known, but also that it is relatively fixed, permitting only slight reconfiguration. If the robot or tool geometry changes, or if the environment is not precisely known, then the robot will not operate properly. An intelligent positioning system based on concepts of eye-hand coordination would offer the flexibility to control a wide variety of robots. An excellent application of such a control approach is in the new modular robot systems where a wide variety of robots can be constructed and modified rapidly with modular components.

In addition to flexibility, this control approach could also be used to obtain high performance from low-cost systems. Currently, substantial engineering and manufacturing costs are incurred to produce a robot with high positional accuracy. These robots tend to be large (10 to 20 times as heavy as their load capacities) and have expensive joint feedback components. An intelligent, vision-based control could provide high accuracy even for low cost existing robots that were designed for master-slave control, or manipulators that have large static deflections.

7.3 Future Work

In the short term, the vision system needs to be implemented to the greatest extent possible with current steam generator robot control systems. The capability of the tracking system should be expanded by upgrading hardware, processing more edge and region data, tracking more than the minimum two tubes, and processing both sets of field data to double the allowable toolhead speed. These improvements would enable the system to be more robust and accommodate higher toolhead speeds if necessary. In addition, the pose determination system will need to be fully integrated with the robot control system so that vision-based robot calibration and position correction feedback can be realized in the field.

In the longer term, the more advanced levels of control described above should be investigated. These ideas need to be evolved into practical technology so that all the benefits enumerated can be realized for complex systems including serial, parallel, portable, and mobile robots.

References

1. Aloimonos, J. Y. and Tsakiris, D. P., 1991, "On The Visual Mathematics of Tracking," *Image and Vision Computing*, Vol. 9, No. 4.
2. B&W Nuclear Technologies, Lynchburg, VA.
3. Ballard, D. H. and Brown, C. M., 1982, *Computer Vision*, Prentice-Hall, Englewood Cliffs, NJ.
4. Berger, M., 1977, *Geometry I and Geometry II*, Springer-Verlag, New York, NY.
5. Brem, L., and Nandhakumar, N., 1993, "A Machine Vision System for Positioning a Teleoperated Servicing Tool Head," *ANS Proceedings of the Fifth Topical Meeting on Robotics and Remote Handling*, Knoxville, TN, Vol. 1.
6. Brem, L., and Nandhakumar, N. , 1994, "A Machine Vision System for Enhancing the Teleoperation of an Industrial Robot," *Machine Vision and Applications*, Vol. 7, No. 3.
7. Boyer, C. B., 1968, *A History of Mathematics*, John Wiley & Sons, New York, NY.
8. Brown, L. B., Merry, J. B., and Wells, D. N., 1986, "Coordinate Measurement With a Tracking Laser Interferometer," *Lasers & Applications*, October.
9. Brown, L. B., and Merry, J. B., 1988, "Interferometric Accuracy in Three-Dimensional Measurements: A Precision Laser Tracking System," *SME Proceedings, Precision Metrology with Coordinate Measurement Systems.*
10. Calkins, J. M., "Real-Time Compensation of Static Deflections in Robotic Manipulators," MS Thesis, VPI&SU, Blacksburg, VA, 1994.
11. Coe, D. H., Fallon, J. B., West, R. L., and Abbott, A. L., 1995, "Finding the Center: Using Incidence to Recover Geometric Features from Single-Image, Monocular Views," *Proceedings of the Europe-China Workshop on Geometrical Modeling and Invariants for Computer Vision.*
12. Corke, P. I., 1993, "High-Performance Visual Servoing for Robot End-Point Control," *SPIE Vol. 2056, Intelligent Robotics and Computer Vision XII.*
13. Craig J. J., 1989, *Introduction to Robotics, Mechanics and Control*, Addison-Wesley, Reading, MA.
14. Driscoll, W. G. and Vaughan, W., 1978, *Handbook of Optics*, McGraw-Hill Book Company, New York, NY.
15. Ealing Electro-Optics, Holliston, MA.
16. Ellis, T., Abbood, A., and Brillault, B., 1992, "Ellipse Detection and Matching with Uncertainty," *Image and Vision Computing*, Vol. 10, No. 5.

17. Fallon, J. B., Reinholtz, C. F., Vallance, R. R., Voruganti, R. S., and Abbott, A. L., 1994a, "Improving Robot System Performance Using Computer Vision," *Proceedings of the 1994 ASME Mechanisms Conference*, Minneapolis, MN.

18. Fallon, J. B., Vallance, R. R., Reinholtz, C. F., and Abbott, A. L., 1994b, "Verifying Robot Toolhead Location in Real Time Using Computer Vision," *Proceedings of the American Nuclear Society annual meeting*, New Orleans, LA.

19. Fallon, J. B., Collier, J. E., Jenkins, A. S., Reinholtz, C. F., and Abbot, A. L., 1995, "Tracking the Position of a Robot with Computer Vision- Mathematical Modeling and Implementation," submitted to the international journal *Mechatronics*.

20. Fu, K. S. Gonzalez, R. C., and Lee, C. S. G., 1987, *Robotics: Control, Sensing, Vision, and Intelligence*, McGraw-Hill Book Company, New York, NY.

21. GAGE-Line Technologies, Inc., Rochester, NY.

22. Ghosh, S. K., 1988, *Analytical Photogrammetry,* Pergamon Books, Elmsford, NY.

23. Glass, S. W., Fallon, J. B., Reinholtz, C. F., Abbott, A. L., and Asher, J. D., 1994a, "Machine Vision Independent Position Verification for a Nuclear Steam Generator Robot," *Proceedings of the Electric Power Research Institute's 13th Nuclear Power Steam Generator Workshop*, La Jolla, CA.

24. Glass, S. W., Fallon, J. B., Reinholtz, C. F., and Abbott, A. L., 1994b, "Machine Vision Calibration for a Nuclear Steam Generator Robot," *Proceedings of the American Nuclear Society Winter Meeting*, Washington, DC.

25. Gonzalez, R. C. and Woods, R. E., 1992, *Digital Image Processing*, Addison-Wesley, Reading, MA.

26. Groover, M. P., Weiss, M., Nagel, R. N., and Odrey, N. G., 1986, *Industrial Robotics: Technology, Programming, and Applications*, McGraw-Hill Book Company, New York, NY.

27. Haralick, R. M., Lee, C. N., Zhuang, X., Vaidya, V. G., and Kim, M. B., 1987, "Pose Estimation from Corresponding Point Data," *Proceedings of the IEEE Computer Society Workshop on Computer Vision.*

28. Haralick, Robert M. and Joo, Hyonam, 1988, "2D-3D Pose Estimation," *Proceedings of the 9th IEEE International Conference on Pattern Recognition.*

29. Haralick, R. M., 1989a, "Monocular Vision Using Inverse Perspective Projection Geometry Analytic Relations," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*

30. Haralick. R. M., 1989b, "Monocular Vision Using Inverse Perspective Projection Geometry Analytic Relations," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*

31. Haralick, R. M., Lee, C. N., Ottenberg, K., and Nolle, M., 1991, "Analysis and Solutions of the Three Point Perspective Pose Estimation Problem," *Proceedings of the 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*

32. Haralick, R. M. and Shapiro, L. G., 1992, *Computer and Robot Vision, Volume I*, Addison-Wesley, Reading, MA.

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

34. Heath, T. L., 1908, *The Thirteen Books of Euclids Elements*, Cambridge University Press, London, England.

35. Herve, G., Sharma, R., Cucka, P., 1991, "Toward Robust Vision-Based Control: Hand/Eye Coordination Without Calibration," *Proceedings of the IEEE International Symposium on Intelligent Control*, Arlington, VA.

36. Horn, B. K. P., 1989, *Robot Vision*, MIT Press, Cambridge, MA.

37. Hung, Y., Yeh, P., and Hardwood, D., 1985, "Passive Ranging To Known Planar Point Sets," *IEEE International Conference on Robotics and Automation*, St. Louis, MO.

38. Image Technology Incorporated, Bedford, MA.

39. Jain, R. C., Martin, W. N., and Aggarwal, J. K., 1979, "Extraction of Moving Object Images Through Change Detection," *Proceedings of the Sixth International Joint Conference on Artificial Intelligence*, Tokyo, Japan.

40. Jain, R. C., 1984, "Segmentation of Frame Sequences Obtained by a Moving Observer," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6.

41. Kabuka, M., McVey, E., Fhironoshita, P., 1988, "An Adaptive Approach for Video Tracking," *IEEE Journal of Robotics and Automation,* Vol. 5, No. 2.

42. Kerninghan, B. W., and Ritchie, D. M., 1988, *The C Programming Language*, Prentice Hall, Englewood Cliffs, NJ.

43. Koivo, A. J., 1989, *Fundamentals for Control of Robotic Manipulators*, John Wiley & Sons, New York, NY.

44. Lenz, R. K., and Tsai, R. Y., 1988, "Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3-D Machine Vision Metrology," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 5.

45. Linnainmaa, S., Hardwood, D., and Davis, L. S., 1988, "Pose Estimation of a Three-Dimensional Object Using Triangle Pairs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 5.

46. Longhurst, R. S., 1967, *Geometrical and Physical Optics*, John Wiley & Sons, New York, NY.

47. Lowe, D. G., 1992, "Robust Model-Based Motion Tracking Through the Integration of Search and Estimation," *International Journal of Computer Vision*, Vol. 8, No. 2.

48. Mair, G. M., 1988, *Industrial Robotics*, Prentice Hall, Englewood Cliffs, NJ.

49. Math Soft Inc., Cambridge, MA.

50. Merritt, E. L., 1949, "Explicit Three-Point Resection in Space," *Photogrammetric Engineering*, Vol. XV, No.4.

51. Merritt, E. L., 1958, *Analytical Photogrammetry*, Pitman Publishing, New York, NY.

52. Mundy, J., and Zisserman, A., 1992, *Geometric Invariance in Computer Vision*, MIT Press, Cambridge, MA.

53. Nussbaum, A., and Phillips, R. A., 1976, *Contemporary Optics for Scientists and Engineers*, Prentice-Hall, Englewood Cliffs, NJ.

54. Ott, R. L., 1993, *An Introduction to Statistical Methods and Data Analysis, Fourth Edition*, Duxbury Press, Belmont, CA.

55. Papanikolopoulos, N. P., and Khosla, P. K., 1993, "Adaptive Robotic Visual Tracking: Theory and Experiments," *IEEE Transactions on Automatic Control,* Vol. 38, No. 3.

56. Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 1992, *Numerical Recipes in C*, Cambridge University Press, Port Chester, NY.

57. Riddle, D. F., 1992, *Analytic Geometry, 5th Edition*, Wadsworth Publishing, Belmont, CA.

58. Robotics and Mechanisms Group, Department of Mechanical Engineering, VPI&SU, 1994, "Vision-Based Control for Robots," Senior Design Project Report, Blacksburg, VA.

59. Rossol, L., 1981, "Vision and Adaptive Robots in General Motors," *Proceedings of the 1st International Conference on Robot vision and Sensory Controls*, IFS Conferences Ltd., Kempston, England.

60. Rothwell, C. A., Zisserman, A., Marinos, C., Forsyth, D. A., and Mundy, J. L., 1992, "Relative Motion and Pose from Arbitrary Plane Curves," *Image and Vision Computing*.

61. Safaee-Rad, R., Smith, K. C., Benhabib, B., and Tchoukanov, I., 1990, "An Analytical Method for the 3D-Location Estimation of Circular Features for an Active-Vision System," *IEEE Proceedings of the International Conference on Systems Man and Cybernetics*.

62. Safaee-Rad, R., Tchoukanov, I., Benhabib, B., and Smith, K. C., 1991, "Accurate Parameter Estimation of Quadratic Curves from Grey-Level Images," *Image Understanding*, Vol. 54, No. 2.

63. Safaee-Rad, R., Smith, K. C., Benhabib, B., and Tchoukanov, I., 1992a, "Constraints on Quadratic Curves Under Perspective Projection," *Image and Vision Computing*.

64. Safaee-Rad, R., Tchoukanov, I., Smith, K. C., and Benhabib, B., 1992b, "Three-Dimensional Location Estimation of Circular Features for Machine Vision," *IEEE Transactions on Robotics and Automation*, Vol. 8, No. 5.

65. Salerno, R. J., Robertshaw, H. H., Horner, C. G., Milsap, W. J., Bennet, D. W., 1992, "Proposed use of a Variable Geometry Truss Manipulator for Radioactive Waste Removal from Underground Storage Tanks," *Proceedings of the 1992 Waste Management Conference*," Tucson, AZ.

66. Salerno, R. J., 1993, *Position Control Strategies for a Modular, Long-Reach, Truss-Type Manipulator*, Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

67. Sandler, Ben-Zion, 1991, *Robotics, Designing the Mechanisms for Automated Machinery*, Prentice Hall, Englewood Cliffs, NJ.

68. Schmidt, R. Q., 1985, "Inspection and Adaptive Robot Applications Based on Three-Dimensional Vision Measurements," *Proceedings of the SPIE- The International Society for Optical Engineering*, Vol. 521.

69. Science Accessories Corporation, Shelton, CT.

70. Shapiro, V., Backalov, I., and Kavardjikov, V., 1995, "Motion Analysis via Interframe Point Correspondence," *Image and Vision Computing*, Vol. 13, No. 2.

71. Shooter, S. B., and Reinholtz, C. F., 1992, "Extrinsic Calibration of Portable Manipulators," *Proceedings of the 22nd ASME Mechanisms Conference*, DE-VOL. 45.

72. Shooter, S. B., Reinholtz, C. F., and Dhande, S. G., 1992, "On The Kinematic Design of Manipulators For Limited Access Workspaces (LAWS)," *Proceedings of the 22nd ASME Mechanisms Conference*, DE-VOL. 45.

73. Silven, O. and Repo, T., 1993, "Experiments with Monocular Visual Tracking and Environment Modeling," *IEEE 4th International Conference on Computer Vision*, Berlin, Germany.

74. Slama, C. C., 1980, *Manual of Photogrammetry*, American Society of Photogrammetry, Falls Church, VA.

75. Slyusarev, G. G., 1984, *Aberration and Optical Design Theory*, Adam Hilger Ltd, Bristol, England.

76. Sony Corporation of America, Park Ridge, NJ.

77. Spatial Positioning Systems, Blacksburg, VA.

78. Tidwell, P. H., Glass, S. W., Hildebrand, J. J., Reinholtz, C. F., Shooter, S. B., 1991, "COBRA- Design and Development of a Manipulator for Nuclear Steam Generator Maintenance," *Proceedings of the 2nd National Applied Mechanisms and Robotics Conference*, November, Vol. 1.

79. Tidwell, P. H., Hendricks, M. W., Fallon, J. B., Stulce, J. R., Beisgen, L. C. V., Reinholtz, C. F., and Maples, A. B., 1993, "Acoustic Pose Determination Applied to an Underwater Mobile Robot," Proceedings of the 3rd National Applied Mechanisms and Robotics Conference, Cincinnati, OH.

80. Tsai, R. Y. and Lenz, R. K., 1988, "Review of RAC-based Camera Calibration," *Vision, Society of Manufacturing Engineers*, November.

81. Tsai, R. Y., 1987, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," *IEEE Journal of Robotics and Automation*, Vol. RA-3, No. 4.

82. Voruganti, R. S., 1995, "Robot System Calibration: Error Identification, Analysis, and Compensation," Ph.D. Dissertation, VPI&SU, Blacksburg, VA.

83. Westmore, D. B., and Wilson, W. J., 1991, "Direct Dynamic Control of a Robot Using an End-Effector Mounted Camera and Kalman Filter Position Estimation," *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, Sacramento, CA.

84. Wolf, P. R., 1974, *Elements of Photogrammetry*, McGraw Hill, New York, NY.

85. Wolfe, W. J., Weber-Sklair, C., Mathis, D., and Magee, M., 1988, "Locating Known Objects in 3-D from a Single Perspective View," *Proceedings of the SPIE, Intelligent Robots and Computer Vision*, Vol. 1002.

86. Wolfe, W. J., Mathis, D., Weber-Sklair, C., and Magee, M., 1991, "The Perspective View of Three Points," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 1.

87. Zeng, X., Mitchell, L. D., Agee, B. L., 1993, "A Laser Position Determination Algorithm for an Automated Mechanical Mobility Measurement System," *Proceedings of the 11th International Modal Analysis Conference*, Vol. 1.

88. Zheng, Q. and Chellappa, R., 1995, "Automatic Feature Point Extraction and Tracking in Image Sequences for Arbitrary Camera Motion," *International Journal of Computer Vision*, Vol. 15, No. 1-2.

Vita

John Barry Fallon was born the sixth eldest of ten children in Pittsburgh, Pennsylvania on February 2, 1963. Following high school, he attended Penn State's University Park campus for one year and then enrolled in Newport News Shipbuilding's Apprentice School. After completing his four-year apprenticeship with honors, Barry worked as a production Supervisor for ten months and then enrolled in the Mechanical Engineering Department at Virginia Tech. While an undergraduate, he participated in Virginia Tech's cooperative education program with Westinghouse Electric Corporation and Newport News Shipbuilding. In 1991, Barry graduated Summa Cum Laude with a B.S. in Mechanical Engineering, a Minor in Mathematics, and a Cooperative Education Certificate. He immediately entered a doctoral program at Virginia Tech, where he conducted robotics research with Professor Charles F. Reinholtz. Barry hopes to contribute his experience, knowledge, and talents to the engineering profession through a variety of roles over a long career.