SCALE ESTIMATION BY A ROBOT IN AN URBAN SEARCH AND RESCUE ENVIRONMENT

A Thesis

by

MAITREYI NANJANATH

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2004

Major Subject: Computer Engineering

SCALE ESTIMATION BY A ROBOT IN AN URBAN SEARCH AND RESCUE ENVIRONMENT

A Thesis

by

MAITREYI NANJANATH

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Approved as to style and content by:

Richard A. Volz (Chair of Committee)	Nancy M. Amato (Member)
Reza Langari (Member)	Valerie M. Taylor (Head of Department)

May 2004

Major Subject: Computer Engineering

ABSTRACT

Scale Estimation by a Robot in an Urban Search and

Rescue Environment. (May 2004)

Maitreyi Nanjanath, B.E., Delhi University

Chair of Advisory Committee: Dr. Richard A. Volz

Urban Search and Rescue (USAR) involves having to enter and explore partially collapsed buildings in search for victims trapped by the collapse. There are many hazards in doing this, because of the possibility of additional collapses, explosions, fires, or flooding of the area being searched. The use of robots for USAR would increase the safety of the operation for the humans involved, and make the operation faster, because the robots could penetrate areas inaccessible to human beings. Teleoperated robots have been deployed in USAR situations to explore confined spaces in the collapsed buildings and send back images of the interior to rescuers. These deployments have resulted in the identification of several problems found during the operation of these robots. This thesis addresses a problem that has been encountered repeatedly in these robots: the determination of the scale of unrecognizable objects in the camera views from the robot. A procedure that would allow the extraction of size using a laser pointer mounted on the robot's camera is described, and an experimental setup and results that verify this procedure have been shown. Finally, ways to extend the procedure have been explored

ACKNOWLEDGMENTS

I sincerely thank Dr. Volz for being my thesis advisor. The extraordinary depth of his knowledge and perception has been a source of inspiration to me. He provided me with guidance and correction whenever I needed it, making several aspects of my work much easier for me. Without his regular feedback at every stage of my thesis, I would not have been able to complete it.

I would like to thank Dr. Amato for her support and encouragement, and for coming especially for my defense though she was on sabbatical at the time. I would like to thank Dr. Langari for agreeing to be on my committee at short notice when Dr. Lee had to leave. I would also like to thank him for having come to my defense though he had been unwell that day, and for his support and encouragement whenever I had problems. I would also like to thank Dr. Lee for having been on my committee and his interest in my work.

I would like to thank Andrew Daugherity for helping with the assembly of my robot. I also thank Ann Hart for providing a great deal of editing and formatting help. She has been instrumental in bringing my thesis to its present form. I would also like to thank Barbara Braby for her assistance in ensuring that my paperwork was complete and in order. I would like to thank Pramod and Yu for their help in solving the problems that came up while writing my thesis.

Finally, I would like to thank my family and my friends for being there for me whenever I needed them.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

Page

LIST OF TABLES

Page

INTRODUCTION

Urban Search and Rescue (USAR) is the rescue of people trapped by disasters involving man-made structures, such as fallen buildings and caved-in trenches [1]. Specially trained personnel are needed to extract these victims, because partially collapsed structures are not safe and may collapse further without notice, trapping or killing would-be rescuers. Fires may start from short circuits in electrical connections or explosions of gas mains. During rescue, roofs often cave in when fire has eaten away at supports, and smoke makes it difficult to breathe. Therefore, buildings have to be strengthened by shoring (adding supports), and the inside of the buildings must be deemed safe before rescuers can enter. Despite all precautions, there is a high likelihood of injury or death among rescuers because of the nature of the task.

Rescue would be made much easier if robots could be introduced to the scene. Since the Oklahoma City bombing in 1995 and the Kobe earthquake in 1996 [2], [3], research has been seeking to solve the problems involved in introducing robots into rescue operations. Robots can fulfill several functions that would increase the speed and effectiveness of rescue operations considerably and make them safer for rescuers. If suitably designed, they can enter smaller places than humans can, and they can locate and bring aid to trapped victims who would otherwise have been missed in the debris because of inaccessibility. Additionally, they can map the interior of buildings so that rescuers get an idea of what to expect as they search the area. The robots can be

---

This thesis follows the style and format of *IEEE Transactions in Robotics and Automation.*

modified to carry and place supplies and lighting at key points within the building. They can be custom-made to the requirements of the rescuers, so that they are unaffected by the hazards in the area.

While introducing robots to USAR is an attractive idea, there are many challenges to be overcome before this can be done effectively. The robots need to be built with safety precautions in mind; they should not themselves pose a hazard to rescuers. They should be able to move effectively through the rubble of a disaster site and not get themselves stuck. They need to be able to localize themselves and not get lost in the enclosed spaces they search (known as voids). They need to have sufficient power supply to complete the search task they have been assigned. Therefore, such robots must be specially designed to meet the needs of search and rescue.

During the search, a robot is sent into a void, and it sends back images of what it views to the operators controlling it. These images need to be interpreted by the rescuers, who then decide if it is necessary for human rescuers to enter that area. If victims are found there, rescuers work to enlarge the entry into the void, shore up the area, and remove the victims. Interpreting the robot images is a challenge; the robot usually moves much closer to the ground, so the images it receives are at a different perspective from what is usual for humans. There is a lot of debris and clutter from the wreckage caused by the disaster, so that objects in the environment are deformed. Consequently, already oddly distorted and broken objects become even more difficult to identify. For example, at the World Trade Center (WTC), a twisted piece of metal

looked exactly like a fireman's boot and could not be identified as a piece of metal until much later, when the robot moved really close to it [3].

Without the usual visual cues, operators have a very difficult time determining what things are. As mentioned in [4], when the scale of an image is unknown, identification poses a real problem. A robot that could estimate and report on the size of an object being viewed in real time would be of great assistance to the rescuers in determining what the object was. This thesis addresses the problem of determining the scale of unrecognizable objects in the camera views from the robot. Scale had not previously been a factor in recognition problems; recognition research focused on correctly identifying the object using features that were often independent of scale. In the case of USAR, however, humans are available to do the identification. The objects to be recognized are often in strange poses, and the surroundings are unfamiliar. Thus, the camera image does not always provide enough information for recognition. We develop a method that provides the teleoperator of the robot with an estimate of the size of the object being viewed, without necessitating major modifications to existing robots being used for USAR.

This thesis investigates the ability of a robot using active vision techniques to determine the size of an object efficiently and effectively, when it is placed at some distance away at an unknown angle with respect to the robot's center of vision. An active vision system is a system that can "manipulate its visual parameters in a controlled manner in order to extract useful data about the scene in time and space"[5]. In this case, the end user is treated as part of the system. He or she moves the robot and

indicates when to acquire an image. The system is active because the images are obtained in accordance with the needs of the system, rather than being fixed and forcing the system to adapt to them. The procedure described here will provide the scale of the object being viewed, the distance of separation of the object from the robot, and its orientation. It relies on the way the image changes as the angle subtended by the camera's central axis changes. Central to the method is the use of a laser pointer mounted parallel to the forward-looking axis of the camera. The laser illuminates a small point on an object being viewed. In order to determine the size and orientation of this object, the teleoperator manipulates the robot to point at two ends of the object. The points illuminated are recorded and used to estimate object size and orientation. The laser pointer also provides the user with a reference point while guiding the robot, thus simplifying the task at hand. Subsequent sections provide background on USAR and the introduction of robots to USAR. Then, we derive equations that determine the scale and orientation and describe the implementation of a small robot vision system that validates this approach. Finally, experimental results on this system are presented, with an analysis of the errors seen, followed by a discussion on how the approach may be improved.

BACKGROUND

Urban Search and Rescue

"Urban search-and-rescue [USAR] involves the location, rescue (extrication), and initial medical stabilization of victims trapped in confined spaces" [6]. Caved-in trenches and collapsed buildings, whether caused by accidents, terrorist attacks, fires or earthquakes, all fall under the category of USAR disasters. There are four operations involved in USAR: Search, rescue, medical support, and technical support.

The first step is the search. Basic search is usually done by civilians who were in the vicinity at the time of the disaster. These untrained volunteers usually find eighty percent of the victims rescued, during the first few hours after the incident[7]. The victims are usually those who were on the periphery of the site and those who were lightly buried but remained conscious and able to call out. Specially trained personnel then make a more detailed search to find victims buried deep in rubble, or in the small confined spaces within the building called voids. This involves entering voids through whatever access is available, and navigating over piles of rubble, crossing fissures and rifts, and moving through restricted spaces. Such areas often have unknown hazards that make rescue operations dangerous and difficult. Often the only access to these voids remains in small air-conditioning ducts, and pipes for water and gas. The trained search may be done either directly by humans; by using search cameras, which are cameras on extensible rods that send back images through the rod; or by using dogs trained to locate humans in the rubble.

The next step is rescue, which is usually done by firefighters or trained and certified rescue personnel. This involves extracting the victims from the area in which they are trapped. Before entering collapsed buildings, the rescuers must perform several preliminary tasks. They must evaluate the structural integrity of the area to determine whether they can safely enter and exit the building and whether the floors they will be entering on will be able to take their load. Shoring may be done to strengthen the walls, and additional roof support may be added. Power, water, and gas supplies must be cut off before rescue operations can begin. The rescuers have to be careful of biological hazards such as decomposing remains and rats that may carry disease. Confined spaces within the building often do not have adequate airflow and may be filled with poisonous or explosive gases. Despite precautions, accidents to rescue workers often happen because of the hazards involved in rescue operations[7].

Usually, the victims are carried out of the void by the rescuers. If they are too deeply trapped, earthmoving equipment is needed to extract them; first the rubble pile over the area is carefully removed, taking care not to dislodge anything that might fall and hurt the victims further, and then the victims are extracted. This often takes hours. Paramedics then provide medical support to these extracted victims, and often to their rescuers also. Finally, technical support is provided by both civilians and paramedics – this includes controlling robots, ensuring food and medical supplies are available, and ensuring that technical devices are ready for use if needed.

In fires, smoke makes it difficult to navigate and see. Timbers often burn through and fall unpredictably, blocking passages that were previously navigable. People

trapped in the burning building are often rendered unconscious by lack of oxygenated air. Extraction of victims involves rescue personnel having to carry backpacks with heavy air cylinders, special fire retardant clothing, masks, and gloves to protect themselves from the heat and smoke of the fire. The weight of this personal protective equipment makes movement difficult and reactions slow. The amount of oxygen that can be carried usually limits an individual's activity to one hour, after which, he or she must return for a new air supply, limiting the efficiency with which the search of a burning building can be carried out.

In the case of earthquakes, access to the damaged area is itself difficult, as happened at the Kobe earthquake[2]. Accidents block roads and cause traffic jams; bridges and flyovers may collapse; power lines get broken and become a danger to rescuers. In this case, the rescuers have to turn off the main power, water, and gas supplies to the area before beginning operations. The work is still dangerous because of the possibility of aftershocks, which can dislodge precariously balanced roofs and buildings, causing renewed damage and destruction. Aftershocks can be disastrous, even though they are usually of less intensity than the original earthquake, as the buildings often sustain preliminary damage in the original earthquake, and finally collapse only during the aftershocks.

Terrorist attacks are the most dangerous USAR situations, because of their unpredictability. Terrorist attacks include bombings, such as happened at Oklahoma City, the crashing of the two airplanes into the WTC twin towers, and possible chemical and biological attacks, for all of which trained rescuers must be prepared.

Introducing robots can make the process of search and rescue safer and faster. The robots can provide preliminary information on the inside of buildings, and help in locating victims. The robots are expendable and can be abandoned in the event of further collapses or problems in retrieving the robot safely. These considerations have created an interest in using robots in urban search and rescue.

Robots in Urban Search and Rescue

Mobile robots are central to robotics-assisted urban search and rescue because they can speed up the process of locating and retrieving victims. The key desired attributes of mobile robots to allow them to do this are [8], [9]:

1. The ability to enter and explore small confined spaces that cannot be accessed by humans or by rescue dogs.
2. The ability to check and monitor different environmental factors, such as temperature and atmospheric gases, continuously. This would allow them to provide early warning of problems that might otherwise go unnoticed by rescue workers until too late.
3. The ability to process images quickly and extensively, often coupled with infrared sensing capabilities, which allow them to do a more complete survey of their surroundings.
4. The ability to monitor the vital signs of a victim and provide first aid or food and water, or both, until the victim can be removed from the area.
5. The ability to carry supplies, such as extra oxygen tanks and tools, for rescuers to use.

6. The ability detect hazardous materials. These robots can be programmed to identify the hazardous material and warn rescuers or take action to neutralize it. Since they can be custom-made to withstand specific chemicals, it would be much safer to deploy them in disasters with chemical or biological hazards, such as toxic gas release, or in defusing bombs.
7. The ability to assess a structure and assist in shoring it for increased stability.

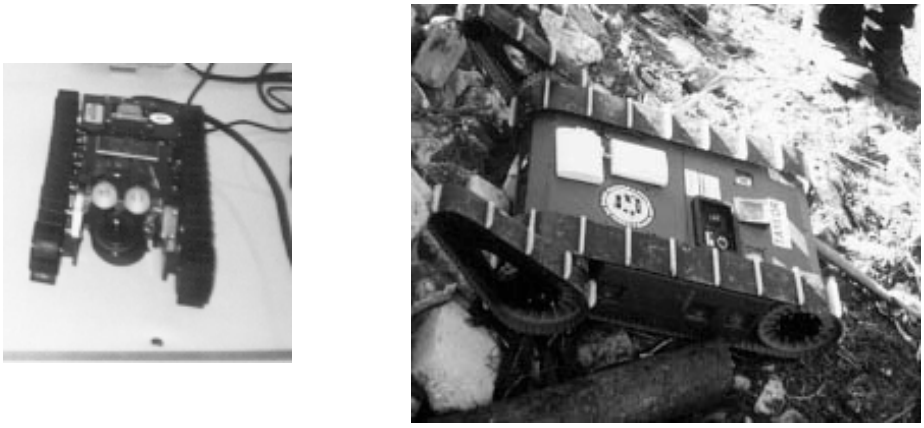

Figure 1: USAR Robots: MicroVGTV(Left) and Urban(Right)

Such robots must be specially designed to meet the needs of search and rescue. With this in view, current research is focusing on how robots may be introduced into USAR environments. The kind of components permissible on the robot (since exposed wires, etc. on the robot may themselves pose safety hazards), and the sensing, communication, and motor capabilities needed for the robot are being studied. Robin Murphy and her group have been studying the needs of USAR personnel and their requirements in these situations [1]. Casper[3] and Micire [10] describe the deployment of teleoperated robots in the WTC USAR situation, and the kind of performance results obtained. The

robots primarily used in these operations were the Urban from RWI, and the MicroVGTV (see Figure 1) from Inuktun. Neither of these robots was specifically designed for USAR; however, they were able to locate victims and provide help in searching voids. Murphy, et al. list several challenges to effective deployment of robots in USAR [1]:

1. Power Supply: The addition of an extra battery to the robot greatly increases its size and makes it unusable for confined spaces. New ways of designing the robot, different battery varieties, or a different method of supplying power is needed to make confined-space-sized robots more practical.
2. Communications: Wireless communication creates a noticeable delay in transmission that is unacceptable in search-and-rescue scenarios. In addition, a lot of wireless dropout (loss of communications) results from fallen debris. While robots communicating through a tether do not have this problem, the tether itself presents snagging problems when it becomes caught on debris along the path. Furthermore, sending images consumes a lot of bandwidth, which means only low-resolution video can be sent effectively within the present power and bandwidth constraints. During a bomb scare, radio frequencies cannot be used for communication as they may trigger the bomb.
3. User Interface: The teleoperated robots load the human operator controlling them with large amounts of information. The operator must process the video feed from the robot, while navigating through the rubble, remembering both

position and directional cues. The interface should have more readily accessible information, without overloading the operator.

4. Hardware: The robots were not manufactured specifically for USAR, so they often did not comply with safety requirements. For example, exposed halogen headlights were used in the Inuktun, which could have cracked, and caused serious problems [3]. The robots should be resistant to heat and waterproof. They should not have exposed wiring that can spark, triggering fires or explosions, and they should be sturdy and able to withstand impact.
5. Sensors: The robots need sensors that can monitor and report their own state. This would relieve the operator of some of the cognitive load, since any problem in movement or damage to the unit would be reported by the robot to the operator, rather than the operator having to retrieve the robot to find out what is wrong. This prevents accidents like the one where the robot impaled itself on a metal rod at WTC [3]. The rod was visible in previous footage but had passed out of sight when the robot actually reached that area, so that the robot became stuck on it.
6. Environment mapping: USAR robots should be able to keep track of the environment already traversed, so that they always know the way back and can also provide a map if needed.
7. Scale: One major point that several rescuers stressed was the lack of information about size in the images sent back by the robot. With little knowledge of the scale of the object being viewed, it became very difficult to

recognize the object for what it was [4]. Three-dimensional range sensors and expensive cameras carry a size and weight penalty and consume more power, which makes them impractical for installation on small rescue robots [4]. Therefore, new methods are needed for estimating the size of an object that are not computationally expensive and do not require addition of heavy equipment to the robots.

8. Artificial Intelligence: Adding artificial intelligence to the robot would be useful in ensuring more complete search coverage; in allowing collaborative teleoperation (to permit a wider area to be covered); and in topological mapping, that is, mapping the environment in terms of obstacles and pathways, rather than in absolute terms. (Topological mapping is more useful for rescuers since it provides information on access routes more directly, and the extra computation needed for accuracy when mapping in absolute terms is not required.) It would also allow for programs such as scripted navigation for stairs, which require complex motion and sensing.

To achieve the above objectives, various different robot forms are being investigated. One avenue of research focuses on marsupial robots for this purpose, while another deals with shape-shifting and serpentine robots, as detailed below.

Current Research in Robots for Urban Search and Rescue

Marsupial and Shape-shifting Robots

Marsupial robots are being considered for USAR for entry into voids where communications and power considerations become problematic. Research on different kinds of marsupial robots is described in [11], [12]. A mother robot would enter the void carrying several smaller daughter robots (Figure 2). The smaller robots would then deploy from it within the void, and would communicate through the mother to send pictures of the area they are traversing. The advantage of this is that the daughter robots can cover a wider area than the mother robot can and return to the mother to be recharged. The overall computational capability can also be increased.

Figure 2: Marsupial Robot

Shape-shifting robots, such as the microVGTV used at WTC can change their shape to fit more easily through small voids and can pass under low roofing.

Reconfigurable Robots and Serpentine Robots

Reconfigurable robots consist of separate, detachable modules that can be reconfigured as needed [13]. These modules can be assembled into walking robots (hexapods), snakes, or wheels. Such robots are useful where the same robot needs to be able to fit through topologically different environments. The snake robot [14] is a specialization of the reconfigurable type in which all the modules are attached end to end. Special devices, such as a camera, may be attached to one end. This form is very versatile, since it can crawl into pipes, raise itself up to view items at higher levels, and does not face the mobility problems most other robots face. The elephantine robot is a modification of the snake, in which one end of the snake is attached to a processing board, giving it better processing capabilities than the snake (Figure 3, [15]).

Figure 3: Serpentine Robot

Specialized Robots

Several specialized robots are being developed, where the robot is geared strongly towards fulfilling one task effectively. The Scout Robot from the University of Minnesota [16] is a small, low-powered reconnaissance robot. It can transmit video using an analog camera and video transmitter, and can be teleoperated. Pipe-crawling robots have been designed at North Carolina State University (NCSU) [17]. These robots are small enough to fit into a six-inch pipe and are made entirely of off-the-shelf components. Robots for fire-fighting are being designed in Japan [18].

Simulation of Search and Rescue Scenarios

Research is also being done in simulating USAR situations. The RoboCup competition [19], which provides a venue for participants to compete with their robots against one another, created a new branch called RoboCup Rescue to foster research and development in USAR [2], [7], [20]. Different levels of difficulty have been set to allow for different robot capabilities. The National Institute of Standards and Technology (NIST) has also created a standard test bed for this purpose [21], [22]. Hopefully, these competitions will help researchers develop improved robot systems that perform better than current ones.

Sensors

An important avenue of research is the improvement of the sensors used in USAR robots. Presently, several different types of sensors are being tested. The artificial nose can sense and identify scents with a high degree of accuracy [23]. The forward-looking

infrared (FLIR) camera provides the same kind of image for an infrared spectrum that a camera would provide for the normal light spectrum [24]. This ability makes it very useful in locating trapped humans using their body heat as an indicator. The omni-directional camera provides an image of all directions at the same time [25]. It has two problems still to be overcome: angular distortion and a missing center portion resulting from the camera axis intersecting that point. Finally, despite having limited angle of vision and low resolution, the ordinary color camera was the most used sensor at the WTC [3].

Software

During the tests carried out at WTC, Dr. Murphy and her group identified several key points that must be addressed for robots to be deployed in USAR successfully [3], [10]. The robots need some image-processing capabilities to process the image data they obtain while searching. They should be able to track their path, and have environment-mapping capabilities. Artificial intelligence is also highly desirable in the robots, so that they could recover from equipment failures or other problems during searches. Finally, some method of providing scale estimates of objects being viewed to the robot operator is needed.

Scale Estimation

At WTC, the researchers found that rescue personnel had considerable difficulty recognizing images sent by the robot. Two major sources of difficulty were the camera point of view on the robot and the lack of visual cues in the damaged buildings. Since

the point of view of the robot was at ground level, images were distorted with respect to human perception, and the humans had to make a change in perception to be able to use the robot effectively. This distortion caused size, shape, and distances to be misjudged. Since there were few recognizable visual cues, the operators did not have familiar objects for comparison to deduce sizes and distances. Operators therefore often made mistakes estimating distances traveled and object sizes, missing several significant remains in the debris[3], that were found in subsequent re-runs of the data collected. [10] and [4] explain how the lack of information on scale caused problems in locating victims at WTC and in other USAR scenarios.

Most scale estimation methods are based on triangulation, where the camera detects the image formed by a light source on an object, and the position of the light source relative to the camera is known. Thus, the camera, light source, and image of the light source form a triangle with some known parameters, from which the unknowns can be calculated. Scale estimation methods using structured light projection have been used for component size measurement in industrial applications[26]-[31]. Structured light projection involves projecting a specific pattern of light onto the object to be measured, so that the camera can retrieve information from the way the light falls on the object being viewed. These systems generally achieve a high accuracy (errors as low as 1 in 20,000) [26]. However, they rely on environment factors being tightly controlled, including assumptions about the background of the object and its texture that do not hold for search and rescue conditions. Also, the projector and the camera usually need

to be placed over a meter apart, which is not feasible for the small size required in USAR robots [32].

Scale has also been researched as a factor in automated object recognition [33], [34]. Most recognition methods, however, mainly focus on working around the problem of identifying the scale of the object, usually by using features that are scale invariant such as color and shape. This is not usable in USAR, because the recognition here is done by humans. The robot needs to provide sufficient information for the human to recognize the object effectively, and scale is an essential part of this information.

SCALE DETERMINATION: APPROACH

We assume we have a robot arm with a camera mounted on the last link. A laser pointer is mounted parallel to the camera lens axis. We move the last link of the robot from a position in which the laser points to one end of the object to a position in which it points to the other end, and acquire images at each end. These images are processed to obtain the distance between the laser point and the image center, and the value of and variation in this distance is used to compute the scale and orientation of the object being measured.

Camera-based measurement methods use cameras that can be modeled on the pinhole concept, i.e., the camera behaves like a pinhole, so that image height is proportional to the object height, and is dependent on the distance of the object from the focal plane. The image height is measurable, so if the focal length of the camera lens and the distance of the object from the lens are known, the object height can be calculated. In our case, this process had to be modified because the camera was not directly analogous to a pinhole camera. Certain distortion effects had to be corrected, before computations could be done. However, as after correction, the camera image may be treated as if it were from a pinhole camera, we describe the basic algorithms first. Distortion correction will be considered in a later section.

Since the actual distance of the object from the lens was unknown, some additional information was needed from the images to compensate for this second unknown. This additional information is provided by the laser pointer. The laser pointer casts a line of

light parallel to the camera axis, which produces a bright point in the camera image. The distance of the bright point from the center of the image can be measured, and can be used to compute the distance of the object from the camera.

We start with some restrictive assumptions, which match the experiment setup used in this thesis, and later in this section extend the derivations to a more general case. Compensatory calculations for deviations from the assumptions we make during these derivations is provided.

Algorithm Development

Restricted Case

We assume that we have a robot arm, with the camera mounted on it such that the central axis of the camera lens (the camera viewing axis) coincides with the last link of the robot. The relationship of the last joint with the laser pointer and the camera axis is shown in Figure 4. We assume the camera viewing axis (z_c) passes through the origin of the last joint's coordinate system (O), such that it is perpendicular to the axis of rotation of the joint (x_ℓ). A laser pointer is mounted on it in the $x_c - z_c$ plane. In order to derive scale and orientation, the last joint is rotated from one end of the object to the other, and images taken at each end. The distance of the laser point from the center of the image in each image is measured, and used to determine size and orientation.

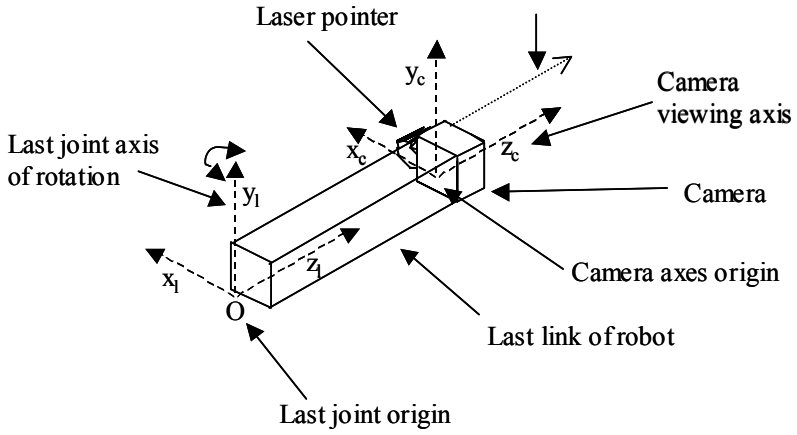

Figure 4: Original Robot-Camera Configuration

In order to derive the relationships for object size and orientation, certain constants related to camera and robot have to be determined. These constants are derived in a calibration step performed before the actual use of the system. This calibration step is described below.

Calibration

For the camera, the actual image height in millimeters must be obtained using the corresponding pixel height. This can be achieved using a multiplier, k, which has units of mm/pixel and is multiplied by pixel height to get the desired height in mm. The scale factor k can be found directly, if the distance between successive CCD elements in the camera in the x and y directions is known. Normally, the camera manufacturing can provide this information. However, in our case, the camera manufacturer was unable to provide us with these CCD element distances. Therefore, k has to be derived during calibration, using the equations described later in this section.

For the robot, two constants need to be determined. These are the distance between the camera axis and laser pointer, denoted by r, and the distance between the camera focal plane and the axis of rotation of the robot, denoted by m (Refer to Figure 5). To determine these constants, we placed a two-dimensional calibration grid perpendicular to the robot base at an unknown distance ℓ_1 from the camera's focal plane. At rest, the camera's central axis was perpendicular to the calibration grid. This is illustrated in Figure 5, where O is the axis about which the last link rotates, OF_1 is the camera axis while the camera is at rest, and B_1 is the center of the camera's focal plane. f is the focal length of the camera, and ℓ_1 is the distance of the focal plane from the calibration grid. Then, d is the distance between laser point and image center in the image and α_0 is the angle subtended by the laser point at the center of the camera's focal plane while the robot is at its rest position. r is the distance between camera axis and laser pointer center, and m is the distance between the axis of rotation of the robot and the camera's focal plane.

Next, the last link is rotated by an angle θ, such that the center of the camera axis moves to the point coinciding with the laser point position on the calibration grid while the robot was at rest. Thus, $OF_2 = OE_1$ now represents the camera axis. r' is the new distance between laser point on the grid and F_2, and is equal to the distance moved by the laser spot across the screen. α_1 is the new angle subtended by the laser point at the center of the focal plane, B_2 is the new center of the camera's focal plane and E_2 is the new location of the laser point on the calibration grid.

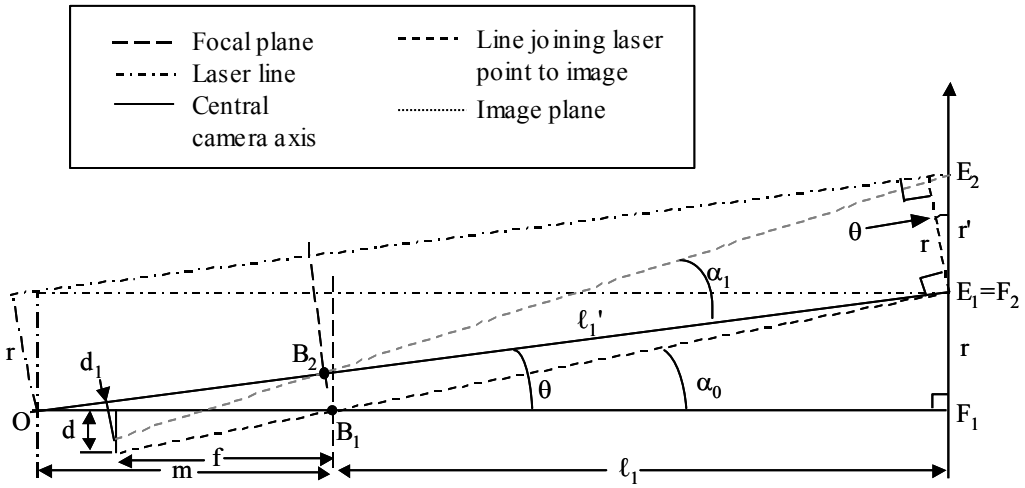

Figure 5: Robot Calibration

We determined r by taking an image of the calibration grid at the rest position of the robot. The grid points visible on the image were used to locate the position on the grid that appeared at the center of the image. This position on the grid would be the point at which the camera axis intersected one of the cross lines on the grid. The horizontal distance between the spot created by the laser pointer and the camera axis intersection as located on the grid is then equal to r. The corresponding distance in the image between image center and laser spot center was labeled d.

In order to obtain the distance between the axis of rotation of the robot and the camera axis (m), we then rotate the last link of the robot by an angle θ about its axis of rotation O, such that the camera center moves to the previously located laser spot center. The distance between the image center and the laser spot in this new image is measured and labeled d_1. Using θ, r, d, d_1, and the focal length f of the camera (usually given by the camera manufacturer), m and k can be derived as follows.

From the figure, we have

$$r' = r / \cos\theta, \tag{1}$$

$$(\ell_1 + m) = r / \tan\theta, \tag{2}$$

In triangle $B_2E_2E_1$, we have the angle at E_1 as

$$\underline{/\ E_1} = \pi/2 + \theta, \tag{3}$$

$$\underline{/\ E_2} = \pi - \alpha_1 - (\pi/2 + \theta) = \pi/2 - \alpha_1 - \theta. \tag{4}$$

Then, using the Law of Sines on triangle $B_2E_2E_1$,

$$r' / \sin\alpha_1 = \ell_1' / \sin(\pi/2 - \alpha_1 - \theta)).$$

Therefore,

$$r' / \sin\alpha_1 = \ell_1' / \cos(\alpha_1 + \theta); \tag{5}$$

also, from triangle OE_1F_1,

$$\ell_1' = (r / \sin\theta) - m. \tag{6}$$

Replacing r' in (5) using (1) and replacing ℓ_1' in (5) using (6), we get:

$$\frac{r}{\cos\theta \cdot \sin\alpha_1} = \frac{(r / \sin\theta) - m}{\cos(\alpha_1 + \theta)}. \tag{7}$$

Therefore,

$$r(\cos(\alpha_1 + \theta)) = ((\ r / \sin\theta\ ) - m\ )\cdot(\cos\theta\cdot\sin\alpha_1).$$

Expanding $\cos(\alpha_1 + \theta)$, we get:

$$r(\cos\alpha_1\cdot\cos\theta - \sin\alpha_1\cdot\sin\theta) = ((\ r / \sin\theta\ ) - m\ )\cdot(\cos\theta\cdot\sin\alpha_1),$$

$$r(\cos\alpha_1\cdot\cos\theta - \sin\alpha_1\cdot\sin\theta) = r(\cos\theta\cdot\sin\alpha_1) / \sin\theta \quad - m\cdot\cos\theta\cdot\sin\alpha_1.$$

Grouping the terms containing r,

$$m\cdot\cos\theta\cdot\sin\alpha_1 = r\cdot\cos\theta\cdot\sin\alpha_1 / \sin\theta - r\cdot\cos\alpha_1\cdot\cos\theta + r\cdot\sin\alpha_1\cdot\sin\theta.$$

Therefore, dividing both sides by $\cos\theta \cdot \sin\alpha_1$,

$$m = r / \sin\theta - r \cdot \cos\alpha_1 / \sin\alpha_1 + r \cdot \sin\theta / \cos\theta,$$

$$m = r / \sin\theta - r / \tan\alpha_1 + r \cdot \tan\theta. \tag{8}$$

Now,

$$\tan\alpha_0 = d \cdot k / f = r / \ell_1, \tag{9}$$

and

$$\tan\alpha_1 = d_1 \cdot k / f.$$

Therefore,

$$\tan\alpha_1 = \frac{d_1 \cdot k}{f} \cdot \frac{r}{\ell_1} \cdot \frac{f}{d \cdot k} = (d_1 \cdot r) / (d \cdot \ell_1). \tag{10}$$

Using (2), (8), (10), we get

$$r / \tan\theta - \ell_1 = r / \sin\theta - d \cdot \ell_1 / d_1 + r \cdot \tan\theta. \tag{11}$$

Therefore,

$$\ell_1 = r / \tan\theta - r / \sin\theta + d \cdot \ell_1 / d_1 - r \cdot \tan\theta,$$

$$\ell_1 - d\ell_1 / d_1 = r / \tan\theta - r / \sin\theta - r \tan\theta.$$

Reversing the signs on both sides and taking ℓ_1 and r common,

$$\ell_1 (d / d_1 - 1) = -r / \tan\theta + r / \sin\theta + r \tan\theta,$$

$$\ell_1 (d / d_1 - 1) = r \left(\frac{-\cos^2\theta}{\sin\theta \cos\theta} + \frac{\cos\theta}{\sin\theta \cos\theta} + \frac{r \sin^2\theta}{\sin\theta \cos\theta} \right).$$

So finally,

$$\ell_1 = r \, (\cos\theta + \sin^2\theta - \cos^2\theta) / ((d/d_1 - 1) \, (\sin\theta \cdot \cos\theta)), \tag{12}$$

and from (2)

$$m = r / \tan\theta - \ell_1, \tag{13}$$

from (9)

$$k = r \cdot f / d \cdot \ell_1. \tag{14}$$

Thus, the value of k provides a conversion from pixel distance to distance in mm. The values obtained here are used in the remaining experiments to calculate the scale of the object being measured.

Scale of the Object

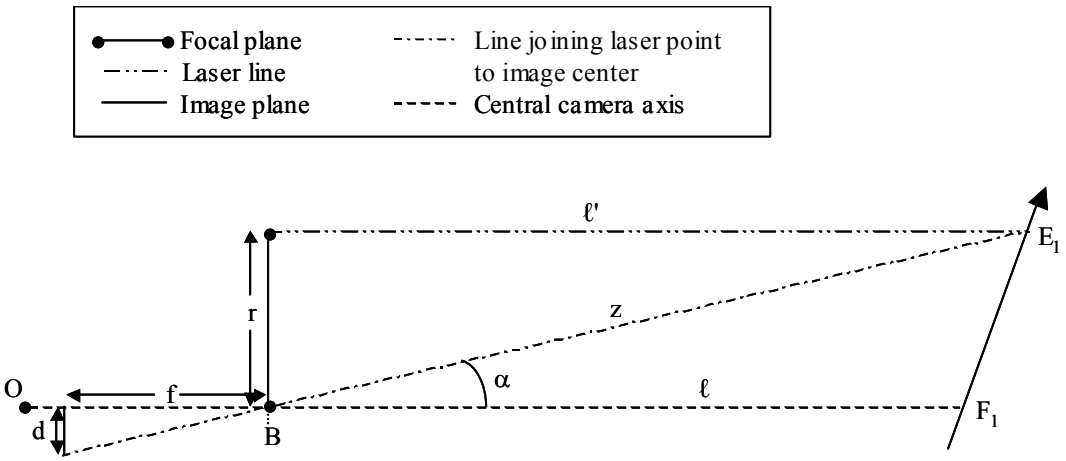

Figure 6: Scale Estimation − Finding z

Let the robot be facing an object placed at an arbitrary orientation. Rotate the last link until the laser pointer spot is at one end of the object (this is equivalent to the point E_1 in Figure 6). Now, the length of the line from the center of the camera's focal plane to the laser spot on the object is represented by z. z can then be computed as

$$z_1 = r / \sin\alpha_1, \tag{15}$$

from Figure 6. Now, Figure 7 shows the system in the same position as Figure 6, except that the image plane and its related terms are not being shown, and the center of rotation of the last link, O, is being shown. Then, the distance of the laser point on the object (E_1) from O is given by x_1 at this position, and the angle subtended by this line (OE_1) is given by γ_1.

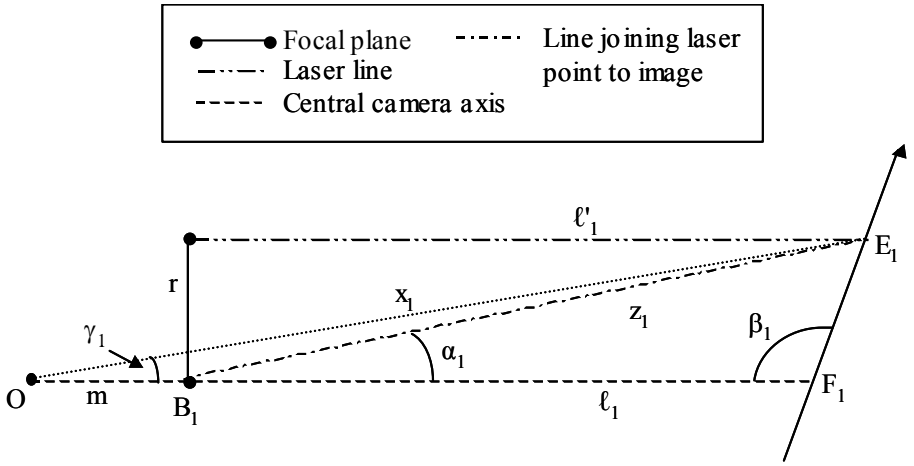

Figure 7: Scale Estimation − Finding x_1 and γ_1

From this figure, using the Law of Cosines, we get:

$$x_1^2 = m^2 + z_1^2 - 2mz_1 \cos(\pi - \alpha_1)). \quad (16)$$

Using the Law of Sines,

$$\sin(\gamma_1) = z_1 \cdot \sin(\pi - \alpha_1) / x_1. \quad (17)$$

Now, let the camera rotate about its axis so that the laser pointer points to the other end of the object, and let the angle of movement be θ. Then, (16) and (17) can be computed similarly for the new position of the camera and laser pointer. Let the values

computed from (16) and (17) for the new position be labeled x_2 and γ_2 respectively (Refer to Figure 8).

Let ρ be the angle between x_1 and x_2. Then,

$$\rho = \gamma_2 + \theta - \gamma_1. \tag{18}$$

Finally, using the cosine rule again, we can find E_2E_3 in Figure 8.

$$E_2E_1 = (x_1^2 + x_2^2 - 2 \cdot x_1 \cdot x_2 \cdot \cos(\rho))^{1/2}, \tag{19}$$

$$\text{Sin}(\varphi) = x_2 \, \text{Sin}(\rho)/E_2E_1, \tag{20}$$

$$\beta_1 = \varphi - \gamma_1. \tag{21}$$

Thus, both the angle of the object with respect to the viewer (β_1) and the length of the object (E_2E_1) can be calculated.

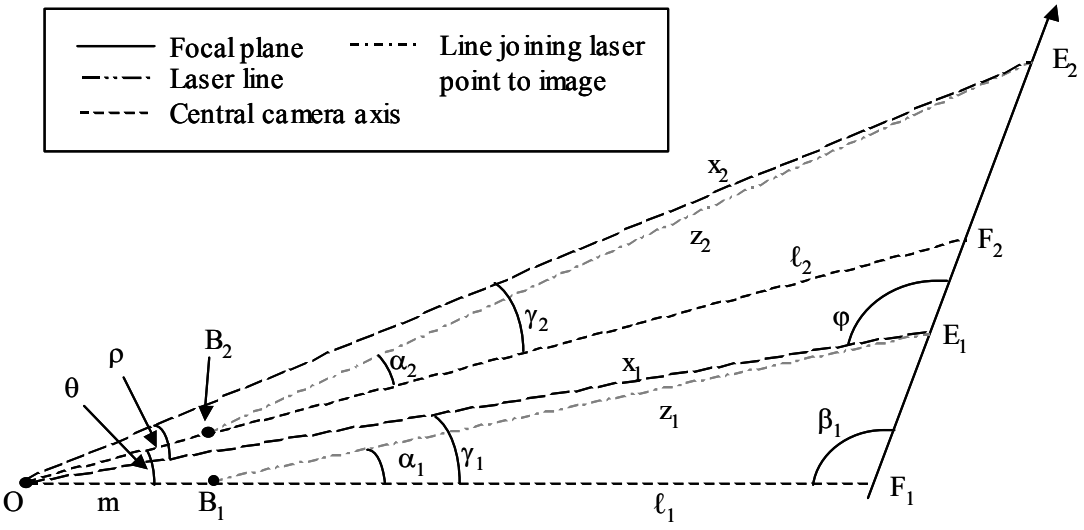

Figure 8: Final Calculations

Detailed Procedure for Obtaining Measurements

Using the derivations and Figure 8 above, the readings taken can be processed in the following manner in determine the size and orientation of the object:

1. Let the readings obtained at one end of the object to be measured be labeled with the subscript '1', and the readings at the other be labeled with the subscript '2'. Let the angle moved by the robot between the two ends of the object be θ.
2. Measurement of the distance between the laser point center on the image and the image center would give the values d_1 and d_2 for readings 1 and 2, respectively.
3. Calculate $\tan\alpha_1$ and $\tan\alpha_2$ from equation 10.
4. Calculate z_1 and z_2 from equation 15 using the values of α obtained.
5. Calculate x_1 and x_2 by taking the square-root of equation 16, using the z's obtained in step 4.
6. Calculate γ_1 and γ_2 from equation 17, using the values of z, α, and x derived in the earlier steps.
7. Calculate ρ using equation 18, from γ_1, γ_2 and θ.
8. Now, put ρ, x_1 and x_2 into equation 19 to get the value of E_2E_1.
9. Finally, find φ from equation 20, and use it to obtain β.

Extension of the Proof: A More General Case

The derivations above assume that the camera is mounted with its axis coincident with the robot's last link and intersecting the last joint's axis of rotation at right angles, and there is a single laser pointer mounted on this camera. This restricts the calculation

of scale and orientation to measurements along a single plane – the plane of rotation used in the calculations. While joints other that the last joint can be used to move the plane itself to different positions, measurements are only valid when both points lie in the plane of rotation of a single joint because the angle of rotation was determined from a single joint sensor. In addition, even if multiple axes of rotation were available, if the laser pointer line is parallel to the plane of motion of the camera axis, the distance between the camera axis and the laser point in the image would never vary. This limitation requires that two laser pointers be mounted on the robot, the second laser pointer being placed so that the plane defined by its axis and the camera viewing axis is perpendicular to the plane of motion of the first camera-pointer system. Therefore, the derivations required to obtain the scale for a more general system which has two laser pointers and is capable of measurements along all directions is described below.

Let the robot be an arbitrary non-Cartesian robot, with several degrees of freedom. Let the camera be mounted on top of the last link of the robot (excluding the gripper, if the robot has one), with the camera viewing axis parallel to the link's z_ℓ axis. The camera coordinate system is chosen so that its z axis (z_c) is along the viewing axis and its x and y axes (x_c and y_c) are in the focal plane. We assume that the rotation of the last joint is about the x_ℓ axis. In addition, we assume that the mounting is such that the x and y axes of the camera and last joint coordinate systems are parallel, and their y-z planes are coincident. Thus, the transformation between these two coordinate systems is simply a translation in the y_ℓ –plane, since the translation required along x_ℓ would be zero. This last condition can be achieved by careful mounting of the camera on the link.

Note that the camera position has been generalized in the sense that the viewing axis (z_c) is no longer required to be coincident with the link z_ℓ axis. Two laser pointers would be mounted, one in the $y_c - z_c$ plane and the other in the $x_c - z_c$ plane. Both laser pointers are mounted so that the laser beam is parallel to the camera z_c axis (except for mounting errors to be calibrated later).

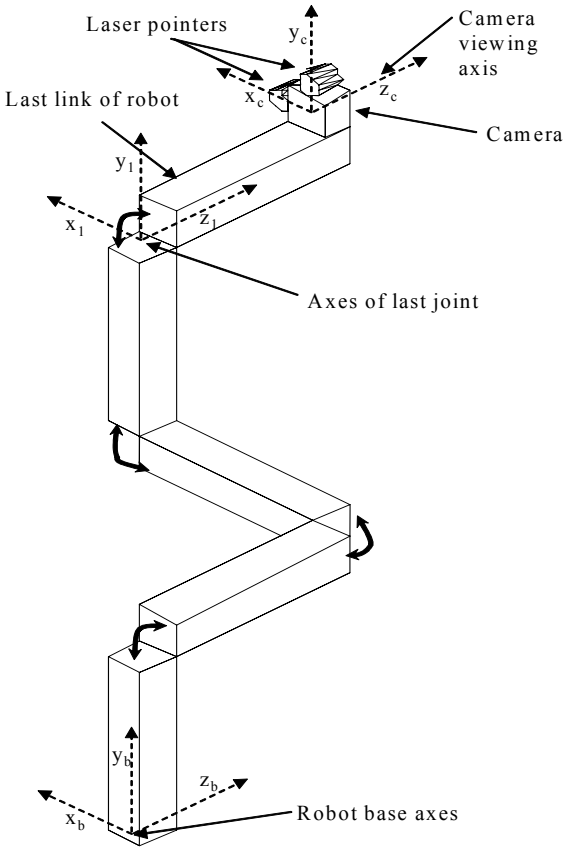

Figure 9: Hypothetical Non-Cartesian Robot

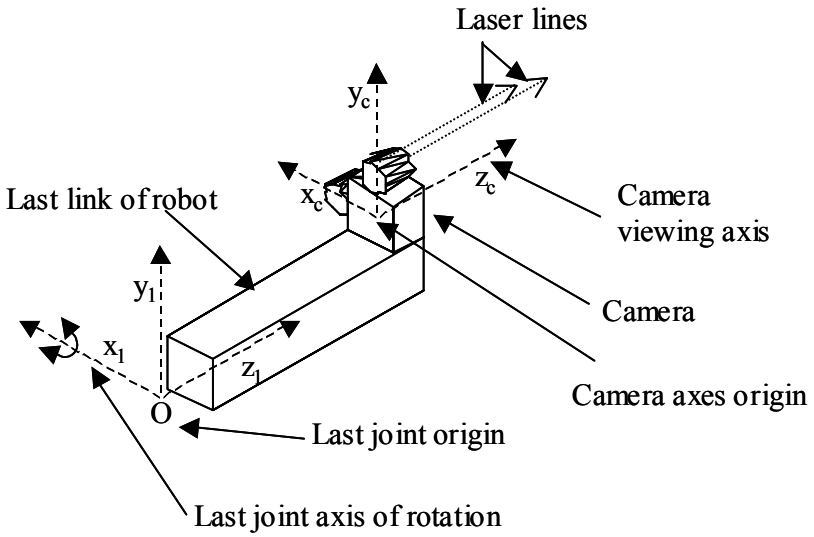

Figure 10: Detailed View of Camera Axes and Last Joint

Such a robot is pictured in Figure 9. The camera axis and its relationship to the last link is shown in Figure 10. Let the robot's axes be defined at the base of the robot. Then, let the x_b-z_b plane through the origin of the robot axes be labeled H_b and the y_b-z_b plane through the origin of the robot axes be labeled V_b. It is assumed that the kinematic relationships of the joints on the robot with the base are known. In that case, relevant points can be projected onto the two planes H_b and V_b. Now, the relevant calculations can be done separately for H_b and V_b.

In order to calculate the scale and orientation of the object, relevant vectors now have to be projected onto the H_b and V_b planes. These projections can be computed from the forward kinematic equations of the robot together with the mounting position and orientation calibration (discussed in the next section). For example, the laser pointer

axis and the camera axis can be projected onto H_b and V_b also, since we make the assumption that they are parallel to the last link.

Two levels of calculations then need to be made: calculations linking the camera position to the mounting link, and calculations linking the mounting link position to the base of the robot. The transformations from the mounting link to the robot base can be represented by a transformation matrix $T_{b\ell}$ that can be derived from the forward kinematic equations of the robot. The camera coordinate system is related to the coordinate system of the last joint by a transformation that can be measured, as seen in Figure 10. After measurements, this transformation can be represented as a transformation matrix $T_{\ell c}$. The section below describes how to calculate this transformation.

Transformation from Camera Coordinate System to Last Link Coordinate System

We need to determine the transformation of the camera coordinate system from the last joint of the robot. As mentioned previously, the camera is assumed to be mounted such that no translation along the x_ℓ axis is required. Therefore, the translation required has to be computed just for the $y_\ell - z_\ell$ plane, which implies that all the required calculations can be done on just this plane. Let the distance of the origin of the coordinate system for the last joint of the robot (O) from the origin of the camera coordinate system (A) be m_{01} (See Figure 11). We assume the intent of the camera and laser mounting are to make their axes parallel to the axis of the last link. Experience has shown that we can do this well enough that at a reasonable distance (1 – 3 feet) the error attributable to misalignment is less than the measurement error of dots projected

on a surface orthogonal to these axes. We perform the camera calibration using a surface with an inscribed grid that is within this distance so that our calibration is limited by the accuracy of the distance measurements. We place a surface with calibrated grid lines in a plane perpendicular to the link axis. The distances of this grid plane from the focal plane and joint origin are unknown. The grid plane is adjusted until its center (the crossing of two grid lines) is at the center of the image.

Let ε be the angle between m_{0l} and the camera viewing axis z_c. This is shown in Figure 11 below. Then, we have to determine m_{01} and η_1.

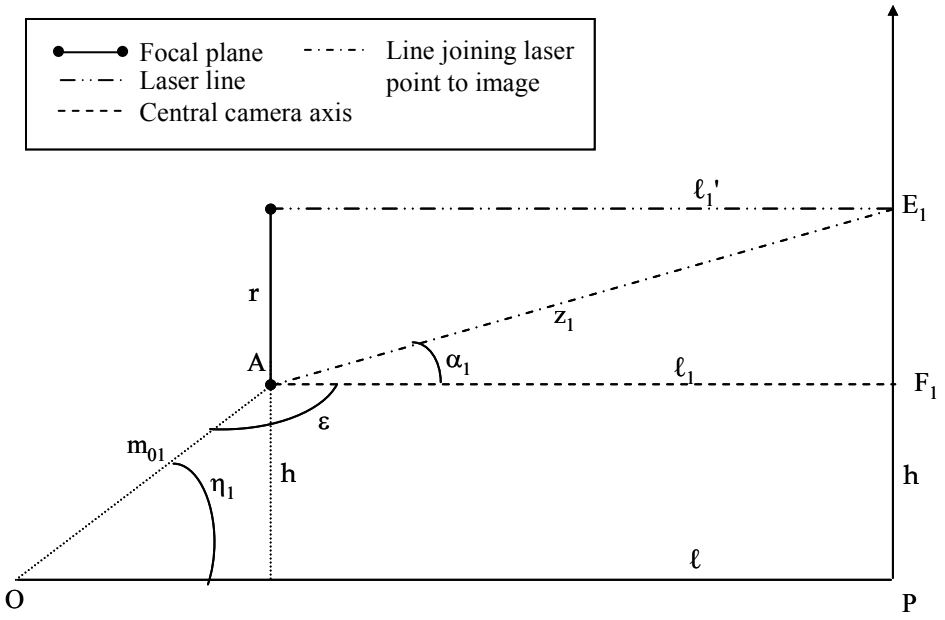

Figure 11: Calibration of Robot - Original Position

Recall that ℓ_1 is the line along the laser pointer and α_1 is the angle from the camera center line to the pointer on the grid intersected by the laser line. The distance E_1F_1 on

the grid is measured. Let r be this measured distance. Since ℓ_1 can be expressed in terms of r and the angle α_1 from (9), we have

$$\ell_1 = r / \tan\alpha_1. \tag{22}$$

Then,

$$h = (\ell - \ell_1) \tan(\eta_1), \tag{23}$$

$$h = m_{01} \sin(\eta_1). \tag{24}$$

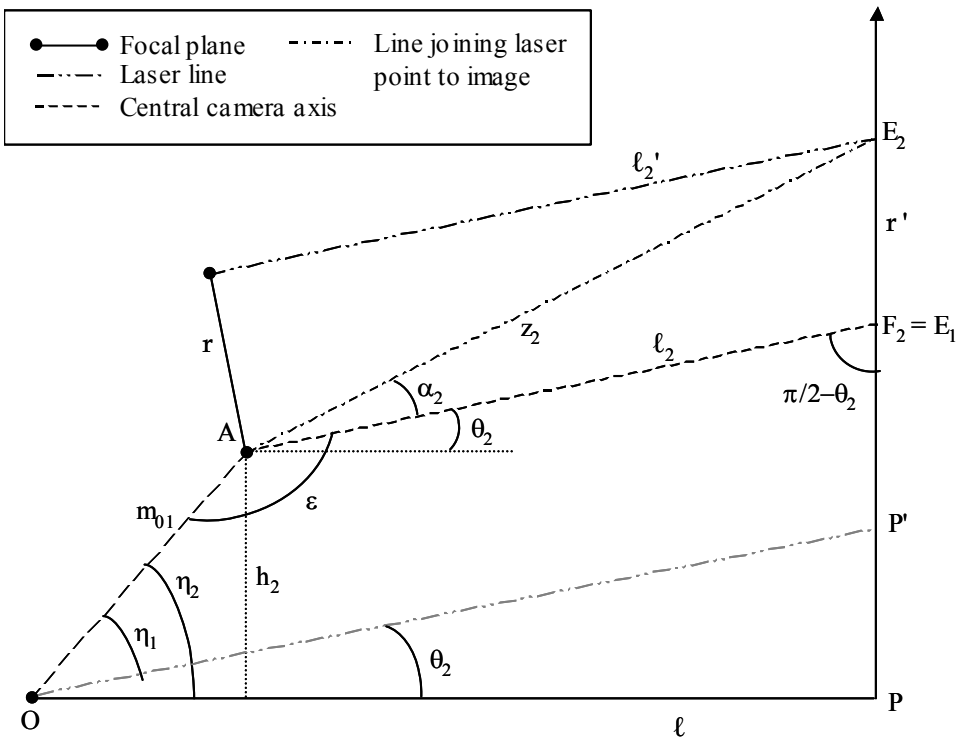

Figure 12: Calibration of Robot: Position After Movement

Then, rotate the last joint through an angle θ_2, such that the new center of the camera image coincides with the previous position of the laser point, as shown in Figure 12. Measure the movement of the laser point on the calibration grid. Let the distance

traveled by the laser point on the grid be r'. Then, at this position we get the following calculations:

$$r' = r/\cos\theta_2, \tag{25}$$

$$\eta_2 = \eta_1 + \theta_2, \tag{26}$$

$$\ell_2' = r/\tan\alpha_2, \tag{27}$$

$$\ell_2 = \ell_2' - r\cdot\tan\theta_2. \tag{28}$$

From (27) and (28),

$$\ell_2 = r/\tan\alpha_2 - r\cdot\tan\theta_2. \tag{29}$$

From Figure 12,

$$\ell - \ell_2\cdot\cos\theta_2 = m_{01}\cdot\cos(\eta_2), \tag{30}$$

$$h_2 = m_{01}\cdot\sin(\eta_2). \tag{31}$$

Now, Figure 12 shows the robot position after the camera axis has been moved over a distance r from the position in Figure 11, on the calibration grid. Therefore, the new intersection point of the camera axis on the calibration grid is given by

$$E_1P = h_2 + \ell_2\sin\theta_2, \tag{32}$$

and from Figure 11,

$$E_1P = h + r. \tag{33}$$

Using (32) and (33), we get

$$h = h_2 - r + \ell_2\cdot\sin\theta_2. \tag{34}$$

From (24), (31) and (34), we can derive m_{01} in terms of η_1, η_2:

$$m_{01}\cdot\sin(\eta_1) = m_{01}\cdot\sin(\eta_2) - r + \ell_2\cdot\sin\theta_2.$$

$$m_{01} = \frac{(r - \ell_2 \sin\theta_2)}{(\sin(\eta_2) - \sin(\eta_1))}. \tag{35}$$

Replacing the equation from (35) in (30), we get ℓ in terms of $\sin(\eta_2)$.

$$\ell - \ell_2 \cos\theta_2 = \frac{\cos(\eta_2)(r - \ell_2 \sin\theta_2)}{(\sin(\eta_2) - \sin(\eta_1))},$$

$$\ell = \ell_2 \cos\theta_2 + \frac{(r - \ell_2 \sin\theta_2)\cos(\eta_2)}{(\sin(\eta_2) - \sin(\eta_1))}. \tag{36}$$

Using h in (23) and (24), we get:

$$m_{01} \sin(\eta_1) = (\ell - \ell_1)\tan(\eta_1),$$

$$m_{01} \cos(\eta_1) = \ell - \ell_1. \tag{37}$$

Using ℓ from (36),

$$m_{01} \cos(\eta_1) = \ell_2 \cos\theta_2 + \frac{(r - \ell_2 \sin\theta_2)\cos(\eta_2)}{(\sin(\eta_2) - \sin(\eta_1))} - \ell_1. \tag{38}$$

Now, replacing m_{01} from (35):

$$\frac{\cos(\eta_1)(r - \ell_2 \sin\theta_2)}{(\sin(\eta_2) - \sin(\eta_1))} = \ell_2 \cos\theta_2 + \frac{(r - \ell_2 \sin\theta_2)\cos(\eta_2)}{(\sin(\eta_2) - \sin(\eta_1))} - \ell_1, \tag{39}$$

$$\frac{(r - \ell_2 \sin\theta_2)(\cos(\eta_1) - \cos(\eta_2))}{(\sin(\eta_2) - \sin(\eta_1))} = \ell_2 \cos\theta_2 - \ell_1,$$

$$\frac{(\cos(\eta_1) - \cos(\eta_2))}{(\sin(\eta_2) - \sin(\eta_1))} = \frac{\ell_2 \cos\theta_2 - \ell_1}{(r - \ell_2 \sin\theta_2)}. \tag{40}$$

Taking just the LHS, and using (26) to remove η_2 from the equation,

$$\text{LHS} = \frac{\cos(\eta_1) - \cos\theta_2 \cos(\eta_1) + \sin\theta_2 \sin(\eta_1)}{\sin\theta_2 \cos(\eta_1) + \cos\theta_2 \sin(\eta_1) - \sin(\eta_1)}.$$

Then, dividing numerator and denominator by $\cos(\eta_1)$,

$$\text{LHS} = \frac{1 - \cos\theta_2 + \sin\theta_2 \tan(\eta_1)}{\sin\theta_2 - (1 - \cos\theta_2)\tan(\eta_1)}.$$

Let

$$(1 - \cos\theta_2) = a,$$

$$\sin\theta_2 = c.$$

Then,

$$\text{LHS} = \frac{a + \tan(\eta_1)c}{c - \tan(\eta_1)a}.$$

Now, all the terms on the RHS are known. So, let

$$\text{RHS} = \frac{\ell_2 \cos\theta_2 - \ell_1}{r - \ell_2 \sin\theta_2} = g.$$

Therefore,

$$\frac{a + \tan(\eta_1)c}{c - \tan(\eta_1)a} = g,$$

$$a + \tan(\eta_1)c = cg - ag\tan(\eta_1),$$

$$(c + ag)\tan(\eta_1) = cg - a,$$

Therefore,

$$\tan(\eta_1) = \frac{cg - a}{(c + ag)}. \tag{41}$$

Now, c and a are known from the measurement of θ_2, and g consists entirely of known quantities, from (22) and (29) Hence, η_1 can be calculated.

Finally, from (35), we get m_{01}, using η_1 from (41) and η_2 from (26).

Therefore, with these calculations, the transformation matrix from the camera coordinate system to the last link's coordinate system can be written as:

$$T_{\ell c} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & m_{01}\sin\eta_1 \\ 0 & 0 & 1 & m_{01}\cos\eta_1 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{42}$$

So far, we have computed the camera displacement from the robot's last link in the $y_\ell - z_\ell$ plane, using the laser pointer mounted in the $y_c - z_c$ plane (call this laser pointer L_y). Now, to use the laser pointer mounted in the $x_c - z_c$ plane (call this laser pointer L_x), we need the corresponding distance of this laser pointer (L_x) from the camera z_c axis. This can be obtained directly from the distance between the center of the camera axis and the spot created by L_x on the calibration grid while the robot is placed such that the camera axis is perpendicular to the calibration grid (which was the original position we had placed the system in, to obtain the transformations previously). We can call this distance r_x. Then, the corresponding distance of L_y (obtained previously) from the axis should be called r_y . At this point, we have all the information required to perform the calculation of the scale and orientation of an arbitrary object.

The forward kinematics of the robot give us the transformation from the coordinate frame of the last link to the coordinate frame of the base of the robot in terms of the set of joint angles. Denote this transformation by $T_{b\ell}(\underline{\theta})$, where $\underline{\theta}$ is the vector of joint angles. Then, given a vector $\underline{w}_\ell$ in the link coordinate system, it can be expressed in terms of the base coordinate system as $w_b = T_{b\ell}(\underline{\theta})\cdot\underline{w}_\ell$. Then, since we have the

transformation from the camera coordinate frame to the last link coordinate frame, we can express the points on the camera image plane in the base coordinate frame as

$$w_b = T_{b\ell}(\underline{\theta}) \cdot T_{\ell c} \cdot w_c, \tag{43}$$

where

$$w_c = [x_c, y_c, -f, 1]^T, \tag{44}$$

and x_c and y_c are the coordinates (in mm) obtained from the camera image.

Procedure for Calculating Scale

Except for certain singular points, either laser pointer alone is sufficient, together with the two readings taken when it points to the respective ends of the object, to determine the scale and orientation. Two lasers are included in the scheme described so that one can be used with the other is at or near a singular point. We thus describe the process in terms of use of just a single laser pointer.

Move the robot so that one of the laser pointers points to one end of the object to be measured. From the joint angles, the aforementioned transformations can be calculated. Next measure the laser point in the image plane. Using the transformations developed above, project each of the following points and lines onto each of the H_b and V_b planes:

- The image point
- The origin of the camera coordinate system (center of the focal plane)
- The z_c axis (the camera viewing axis)
- The line of the laser beam (which is parallel to the z_c axis)

The projections of the z_c axis and the laser line will still be parallel because projections of parallel lines are parallel. The projections of the image point (D_P), the

actual point on the object (E_P) and the center of focal plane (O_{cP}) will lie on a straight line because projections of straight lines are straight lines. Thus, we will have the geometry of Figure 13.

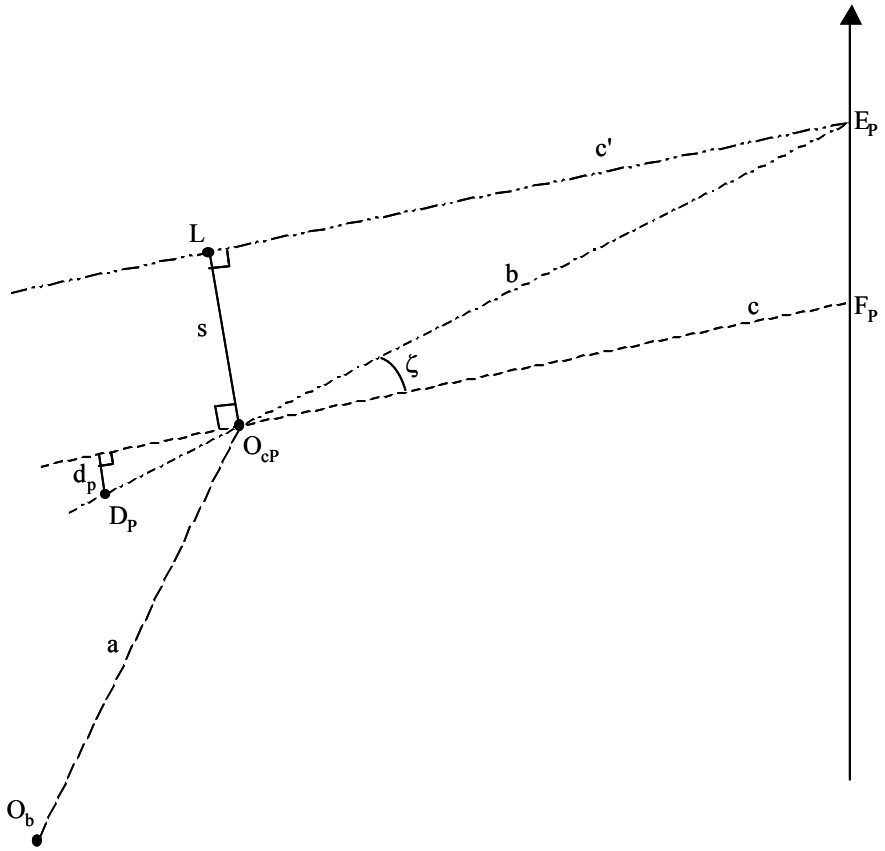

Figure 13: Calculating Scale for 3 Dimensions

The projection of the camera viewing axis z_c and the laser beam lines can be done by choosing two points on each line, projecting the points onto H_b and V_b using the transformations discussed previously, and then finding the equation of the projected line by joining the two projected points. This is detailed in what follows.

In the camera coordinate system, z_c can be represented by the line:

$$\mathbf{z_c} = [\ 0,\ 0,\ z,\ 1]^T, \tag{45}$$

where z is variable. Therefore, any two points chosen on this line, say $\underline{z}_{c1}$ and $\underline{z}_{c2}$ would be represented in the camera coordinate system as

$$\underline{z}_{c1} = [0,\ 0,\ z_1,\ 1]^T \quad \text{and} \quad \underline{z}_{c2} = [0,\ 0,\ z_2,\ 1]^T. \tag{46}$$

Then, let the points in the base coordinate system be $\underline{z}_{b1}$ and $\underline{z}_{b2}$. We have:

$$\underline{z}_{b1} = T_{b\ell}(\underline{\theta}) \cdot T_{\ell c} \cdot \underline{z}_{c1} = [\ x_{zb1},\ y_{zb1},\ z_{zb1},\ 1\ ]^T, \tag{47}$$

and

$$\underline{z}_{b2} = T_{b\ell}(\underline{\theta}) \cdot T_{\ell c} \cdot \underline{z}_{c2} = [\ x_{zb2},\ y_{zb2},\ z_{zb2},\ 1\ ]^T. \tag{48}$$

Since H_b is the $x_b - z_b$ plane, and V_b is the $y_b - z_b$ plane, the equivalent points on these two planes are

$$\underline{z}_{b1}{}^H = (x_{zb1},\ z_{zb1}), \qquad \underline{z}_{b1}{}^V = (y_{zb1},\ z_{zb1}), \tag{49}$$

and

$$\underline{z}_{b2}{}^H = (x_{zb2},\ z_{zb2}), \qquad \underline{z}_{b2}{}^V = (y_{zb2},\ z_{zb2}). \tag{50}$$

Thus, these points can be formed into a line in each plane. The line in the H_b plane would be given by:

$$(x - x_{zb1}) / (z - z_{zb1}) = (x_{zb2} - x_{zb1}) / (z_{zb2} - z_{zb1}). \tag{51}$$

Similarly, the laser beam lines for laser pointers L_x and L_y can be represented in the camera coordinate system as

$$L_x = [\ r_x,\ 0,\ z,\ 1]^T, \tag{52}$$

$$L_y = [\ 0,\ r_y,\ z,\ 1]^T, \tag{53}$$

where z is variable. These can also be projected onto the planes H_b and V_b of the base coordinate system, in a similar manner as described above for the line z_c. Then, the

perpendicular distance between the projected z_c line and the projected laser beam line for each laser beam needs to be computed. This then is the distance s shown in Figure 13. The image point D can be represented by

$$D = [x_{dc}, y_{dc}, -f, 1]^T. \quad (54)$$

Once D has been projected onto the projection planes (call the projected point D_P), the length of the perpendicular (in mm) from D to the projected camera axis needs to be computed. This length is shown in Figure 13 as d_p. Then, ζ can be calculated for the projection from the equation:

$$\tan\zeta = d_p / f. \quad (55)$$

With s and ζ known, the value of $O_{cP}E_P$ (b) and its orientation with respect to O_b can be calculated. Therefore, the coordinates of E_P can be computed on both the projection planes using these computations. Note that the H_b projection plane gives the x and z coordinates for E, and the V_b plane gives the y and z coordinates for the plane. The z coordinates should be equal; the magnitude of the disagreement in the z coordinates can be used as a measure of the error in the calculations for the corresponding laser pointer. A very large difference may mean that laser pointer is close to a singular point, and the computations based on its readings should not be used. For general purposes however, when the agreement between the two z values is good, the average of the two can be taken as the required z.

Now, let the calculation of E_P from the first reading be labeled E_{P1}. Move the robot so that the same laser pointer used earlier now points to the other end of the object.

Then, repeat the projections and calculations above, to get the point E_{P2}. Now, the actual distance between the two points measured is given by $E_{P1}E_{P2}$:

If $E_{P1} = [x_{E1}, y_{E1}, z_{E1}, 1]^T$ and $E_{P2} = [x_{E2}, y_{E2}, z_{E2}, 1]^T$ then the length

$$E_{P1}E_{P2} = ((x_{E1} - x_{E2})^2 + (y_{E1} - y_{E2})^2 + (z_{E1} - z_{E2})^2)^{1/2}, \quad (56)$$

and the orientation required is the orientation of the line $\mathbf{E_{P1}E_{P2}}$ formed by these two points, relative to the coordinate system of the base.

Error Due to Laser Misalignment

There may be an alignment problem where the laser pointer mounted on the robot could not be made perfectly parallel to the camera's central axis. While this would not show up at close range, the error caused by this would become very visible in the case of long-range measurements. Since the error in the alignment of each laser pointer is independent of the other, the error factor needs to be computed separately for each one. However, the calculations are the same for both the laser pointers, so we show the error factor correction required for the laser pointer L_x below.

The misalignment error in L_x can consist of a tilt of the laser pointer towards z_c which can be taken to be in the $x_c - z_c$ plane, and a twisting error, where the laser pointer is tilted in the $y_c - z_c$ plane. These two components can be computed separately. While performing the computations described in the previous section, the error component in the $x_c - z_c$ plane needs to be used for corrections in the projection onto the H_b plane, and the error component in the $y_c - z_c$ plane needs to be used for corrections in the projection onto the V_b plane. The computations required for this compensation

are therefore derived for the projection onto one of the two planes; the error in the second plane can be calculated using the same equations.

The alignment error component in one plane is shown in Figure 14. Let the angle of the laser pointer with respect to the camera z_c axis be ϕ, taking a misalignment angle that makes the laser beam tilt towards the axis z_c as being positive. Then, take the effective height of the laser spot to be r'. We can derive r' in terms of r as follows:

$$r' = \ell.\tan\alpha, \tag{57}$$

$$r = \ell \tan\alpha + \ell \tan\phi. \tag{58}$$

Therefore, from (57) and (58),

$$\ell = r / (\tan\alpha + \tan\phi), \tag{59}$$

Which gives-

$$r' = (r \tan\alpha) / (\tan\alpha + \tan\phi). \tag{60}$$

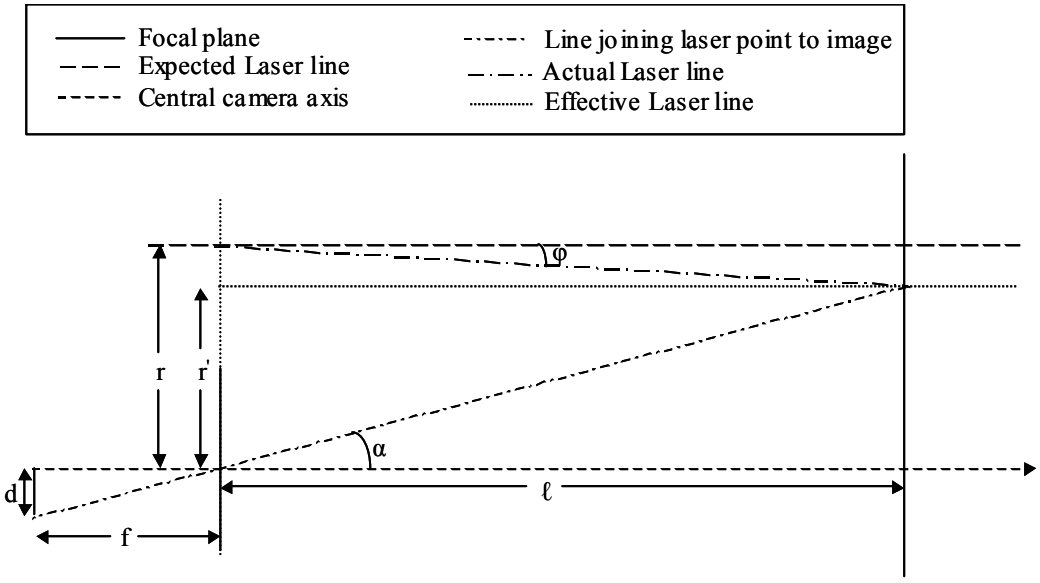

Figure 14: Laser Misalignment

When the laser pointer falls below the camera z_c axis, which would happen if the object is far enough away, then the value of tanα becomes negative. Similarly, if the laser were angled away from the camera z_c axis, tanϕ becomes negative.

Tanϕ can be determined by mounting a calibration grid perpendicular to the last link of the robot. Then, note the position of the laser spot center on the calibration grid. Now, move the calibration grid back by a measured amount, and note the change in the position of the laser spot center. The horizontal difference between the original position of the spot and the new position of the spot on the grid, divided by the distance by which the grid was moved becomes the value of tanϕ in the $x_c - z_c$ plane. Similarly, the vertical difference in the distance becomes the value of tanϕ in the $y_c - z_c$ plane. This misalignment compensation needs to be applied to the laser beam line while performing the projection onto H_b and V_b, by shifting the line so that r' is used instead of r as the distance from the camera axis in the projection. Then, the equation of the line for L_x would become:

$$L_x = [\, r'_x,\ 0,\ z,\ 1]^T, \tag{61}$$

where r'_x is derived from equation (60) by using r_x in place of r, and the corresponding α and ϕ values. This compensation needs to be done separately for each of the laser pointers.

Distortion Removal

Wide-angle cameras are likely to be used for USAR since they provide a wider field of view than normal cameras do. The wide-angle lens used in such cameras results in a

highly visible distortion error in the image produced. The center area of the image remains relatively undistorted, but the edges show a pronounced barrel distortion [35]. Barrel distortion is a form of cylindrical distortion, i.e. the distortion is symmetric about the center of the image, that causes the edges of the image to appear curved in, so that the center seems to bulge outwards (see Figure 15). [36] describes a method of calibrating an off-the-shelf camera mounted on a robot, and removing the distortion in the image, based on the assumption that there is only cylindrical distortion (barrel or pincushion distortion, which are symmetric about the center of the image) present. The model for this distortion is represented as:

$$D_x = X_d (\kappa_1 r^2 + \kappa_2 r^4 + \ldots), \tag{62}$$

$$D_y = Y_d (\kappa_1 r^2 + \kappa_2 r^4 + \ldots), \tag{63}$$

where

$$r = (X_d^2 + Y_d^2)^{1/2}; \tag{64}$$

κ_1 and κ_2 are the lens distortion coefficients and need to be found;

D_x and D_y are the values of the distortion along x and y respectively;

and X_d and Y_d are the measured coordinate positions of the distorted pixels with the origin at the center of the image.

These equations are based on the lens distortion modeling done in [37], and simplified by Tsai in [36]. The attempt to model the distortion using these equations for the camera used in my setup, however, was unsuccessful. A detailed description of the attempt, with calculations, is provided in the appendix.

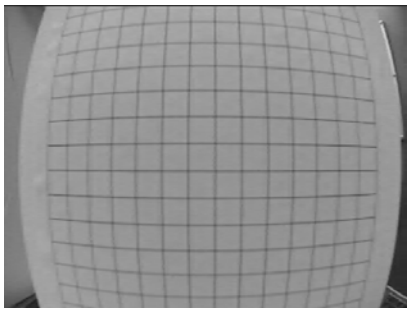

Figure 15: Barrel Distortion

Instead, a more direct distortion-removal method, that would be valid for all cameras, was used. A two-dimensional calibration grid was mounted perpendicular to the camera viewing axis. The grid consisted of regularly spaced horizontal and vertical lines on a flat board, and the spacing between the lines was known. A picture of this grid was taken and analyzed to locate the cross-points where the horizontal and vertical lines intersected. Then, the undistorted center of the image was analyzed. The distance between the cross-points at the center was taken to be the desired distance for the rest of the image. Using this value, the desired position of each point in the rest of the image was computed using linear interpolation. Then, the distorted image was also interpolated to get a set of actual versus distorted image positions. This formed a mesh as shown in Figure 16.

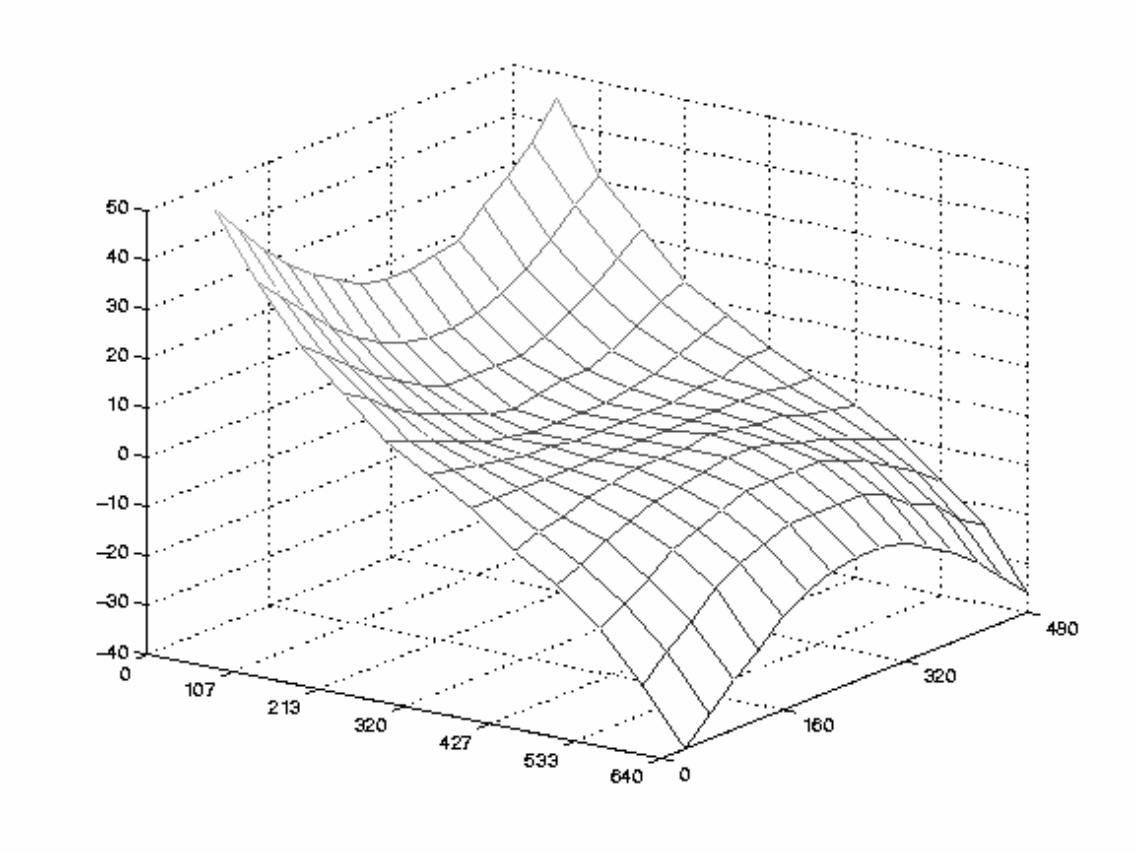

Figure 16: Distortion Mesh

The curvature of the mesh increases as you move away from the center. The height of any point on the mesh represents the amount of distortion at that point Therefore, the corners of the image show the most distortion. Also, this mesh shows that certain points on the original image have to be mapped to more than one point on the distortion-corrected image, due to the ideal image getting ‘compressed’ by the wide angle lens, resulting in blurring of the image. Also, as we move towards the edge, the correction becomes less and less accurate. The desired image was computed using this mesh, by taking every pixel of the mesh and computing the corresponding points on the new

image and the original for that pixel. This mapping was then saved, and used for all the experiment images taken. An assumption being made here is that the distortion does not vary with distance of the image from the camera; it only varies with the relative position of the pixel. This can be taken to be true from the various calibration techniques described in [36], [38]-[48].

The resulting image was much larger than the original, since the border pixels had all expanded. The image was then trimmed to fit 640 x 480 rectangle in size, along all four edges, keeping the center in place, because the central x axis and central y axis usually showed no expansion, so that the original distortion-corrected image had concave edges and acute-angled corners. Trimming the image made it computationally simpler to use and did not affect the results. The resultant distortion-corrected and edge-trimmed image for Figure 15 is shown in Figure 17. Some distortion is visible along the edges, but as will be shown in the calculations, these do not affect the computations because the portion of the image important for computation is the center area, which has had its distortion corrected.

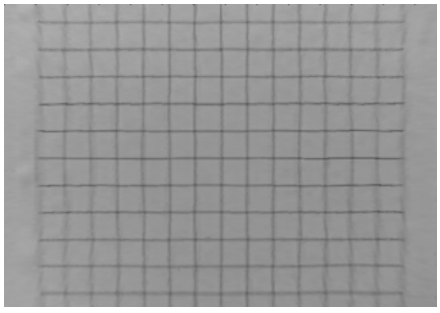

Figure 17: Distortion-Corrected Image

For comparison, a second image and its distortion-corrected equivalent are shown in Figure 18.

Figure 18: Uncorrected Image (Left) and Corrected Image (Right)

EXPERIMENT SETUP

Experiment Description

The experiments performed are intended to verify the applicability and correctness of the mathematical derivations developed in the previous section, and to identify sources of error and the corrections needed for these errors. These experiments were performed using a robot arm with two degrees of freedom, on the tip of which a camera and laser pointer are mounted.

Two rounds of experiments were done. The first round involved short-range measurements, where the camera was placed close to the screen. The measurement ranges were correspondingly shorter. The second involved long-range experiments, with the camera at a distance of over 5 feet.

The first round was intended to prove the general principles of the process of scale estimation. This consisted of five sets of experiments. The robot was mounted just beyond one end of a Lego runway. A ruler was mounted on a screen, and the screen attached to a platform made of Lego bricks. This platform could then be attached at different positions on the runway, with the advantage that the distance of the platform from the robot could be measured easily, and from this the distance of the screen from the robot could be derived.

The second round was intended to examine the effects of scaling up the environment in which the setup has to function. In these experiments, the robot was placed on a table facing a wall at a distance of over 6 feet away. Markings were placed

on the wall at equal intervals (200 mm apart). The robot was then commanded to move from one marking to the other, and measure the separation between the markings.

The results of the preliminary calibration and these experiments are described in the next section.

Equipment

Robot Description

The robot used in these experiments consists of a base, upper arm and forearm. The robot arm was purchased from LynxMotion. LynxMotion provides assembly kits for robot hobbyists and students, so the robots are affordable and easily modifiable, which makes them ideal for the experiments. The arms are made of laser-cut Lexan [49], [50]. This material is light and sturdy, so the final assembly is a lightweight robot with most of the weight being the motors at the joints.

A camera and a laser pointer are mounted on the forearm. Figure 19 shows the setup of the robot. As shown in Figure 20, the robot arm assembly consists of three motors, the base motor providing the pitch (motion about the y axis) and the other two providing roll. Since the axes of motors 2 and 3 are in the same plane, only two degrees of freedom are obtained, which is a constraint on the scope of the experiments.

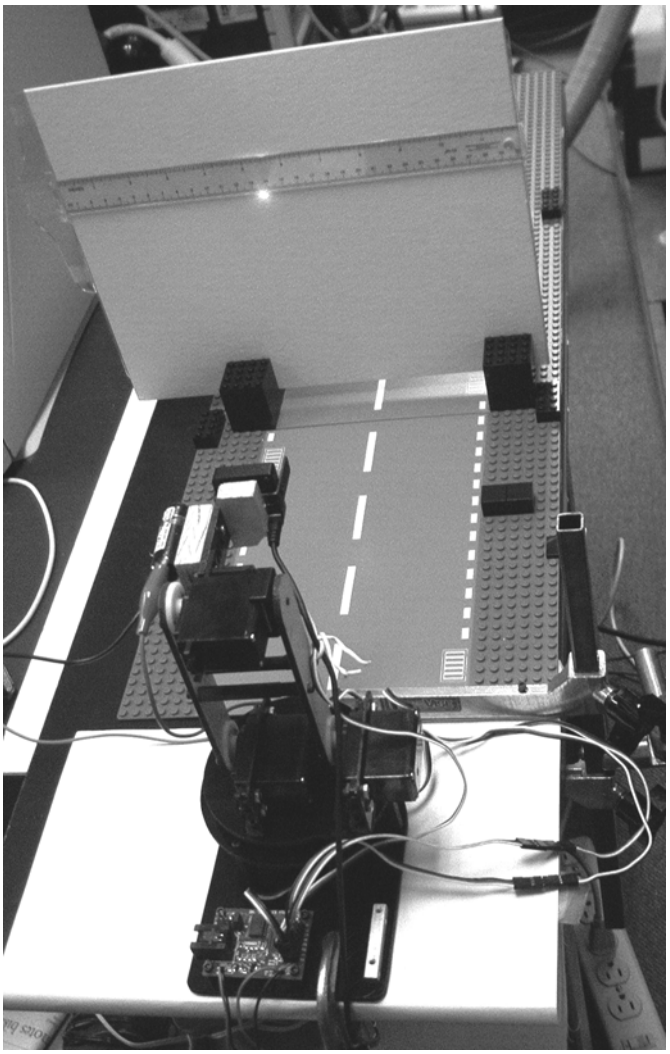

Figure 19: Experiment Setup

The base, shoulder, and forearm assembly was completed, and a camera and laser pointer were mounted in place of the gripper that came with the assembly kit. The camera was mounted on the forearm so that its body was normal to the arm; at rest, the camera viewing axis was horizontal to the ground. The laser pointer is mounted on the forearm to the side of the camera, and parallel to the camera viewing axis.

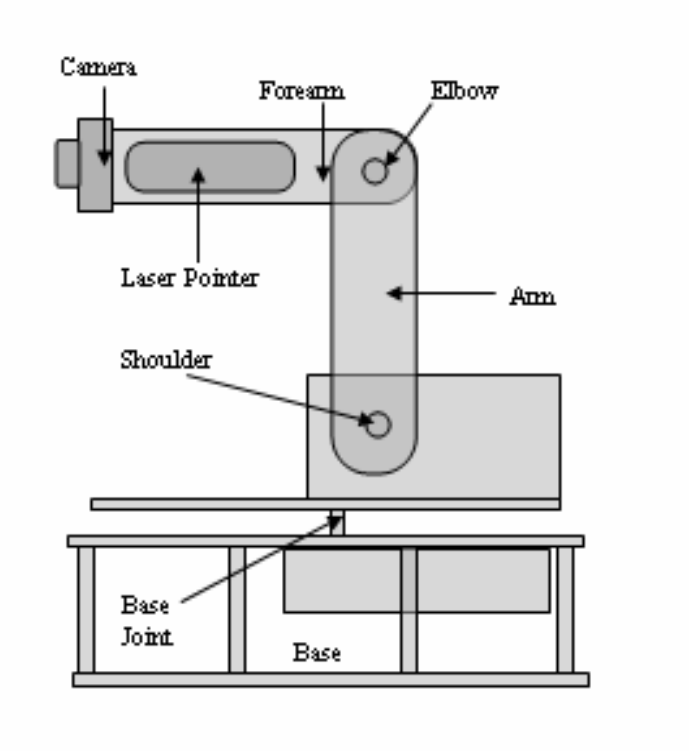

Figure 20: Robot Arm Schematic
Each joint provides one degree of freedom.

The robot uses servomotors (Hitec HS-422) for controlling movement. A servomotor is a motor that comes with control circuits and a potentiometer, which allow the motor to turn through a specified angle and stop. The motor takes a pulse as input and moves to the angle specified by the width of the pulse. Normally, a servomotor can be moved through a 180-degree angle of rotation. A 1ms pulse moves the motor by an angle of 90 degrees. The servomotor controls motor movement proportionally; that is, more power is used when turning through a greater angle.

The servomotors are controlled through a Mini SSC II (Serial Servo Controller). This takes a 8-bit value as input and uses it to drive the motor through the required angle. A detailed description of how the SSC works can be found in [51]. The SSC can

control the servomotors through 180° of motion, giving a 0.72° resolution. The SSC was connected to a PC through the serial port, using a modular line with a DB-modular adapter.

Each motor was then directly controlled using a Java program written for that purpose.

Camera Description

The camera used is a CC-7 HAD CCD camera from CD3 Security (St. Louis, MO; [52]),which provides small surveillance cameras. The camera is only 29mm x 29mm in size and is very light. It connects with the computer through a TV-video-to-USB cable adapter. It generates images with resolution up to 640 x 480 pixels, 24-bit color. The camera uses a wide-angle lens, which shows some distortion at the edges of the image that must be corrected before the image can be used for measurements. The camera is based on [53] which describes the design of a ‘pinhole CCD’ wide-angle zoom lens.

Laser Pointer

The laser pointer is a small, keychain pointer sold for use in classrooms. It was mounted on the camera by removing its batteries, drilling holes in the battery compartment and using these holes to attach it to the robot arm. It was then wired to an external battery. The laser point generated shows some spread, so the center of the point must be located in each image taken. In these experiments, we located the point by hand. Since the laser point is the brightest point in the image, a relatively simple algorithm could be designed to locate the point, also.

RESULTS

Calibration Results

The r (separation of camera from laser pointer), d (separation of image center from laser point at the perpendicular screen position), d_1 (separation of image center from laser point when the image center is at the previous laser point center position) and θ (angle moved to reach previous laser point center from the original position) were measured as described in the calibration section of the previous section. The focal length (f) was given by the manufacturer. The width of the squares of the calibration grid was 14 mm. Using the data (also shown in TABLE I), as given below, the values of m, ℓ_1, and k were derived:

$r = 28.667$

$d = 90.833$

$d_1 = 88.833$

$\theta = 0.0837758$

$f = 3.7$

Therefore:

$\ell_1 = r(\cos\theta + \sin^2\theta - \cos^2\theta) / (d/d_1 - 1)(\sin\theta \cos\theta)$

$= 160.290199$

$m = r/\tan\theta - \ell_1$

$= 28.667/0.083972342 - 160.290199$

$= 181.096016$

$k = rf/d\ell_1$

$= 0.007285$

TABLE I

CALIBRATION MEASUREMENTS

Rdg. #	d	d_1	r	Motor angle μ	$\theta = \mu * \pi / 250$
1	91	88.5	30	127 – 121 = 6	0.0753982
2	91.5	89.5	28	127 – 120 = 7	0.0879646
3	90	88.5	28	127 – 120 = 7	0.0879646
Avg	90.833	88.833	28.667	6.667	0.0837758

Short Range Experiments

In the short-range experiments, a ruler was mounted on a screen, and the camera was moved across selected portions of it taking readings. The measurement results are shown in the table (TABLE II) below.

TABLE II

SHORT RANGE RESULTS

Set #	# of meas.	Screen distance (mm)	largest meas. (mm)	smallest meas. (mm)	Max d (pixels)	Min d (pixels)	Avg. Error (%)	Std. Dev.(%)
1	5	386	50.8	44.0	77	70	0.16	3.18
2	13	378	79.4	30.0	85	68	2.05	9.91
3	12	466	54.0	47.6	52	46	-1.71	4.95
4	7	658	54.0	25.4	28	25	1.05	10.96
5	8	850	60.0	24.0	17	17	3.61	4.90

As can be seen in the results, the readings were very evenly distributed. The error showed a slight increase as the distance from the camera increased. In the first set of readings, one invalid data point was obtained. These readings are shown below (TABLE III).

TABLE III

RULER DISTANCE RESULTS FOR DATA SET 1

Motor angle 1	Angle 1 (radians)	d (pix.)	Motor angle 2	Angle 2 (radians)	d (pix.)	Meas. dist. (mm)	Actual dist. (mm)	Error	Error %
99	-0.35186	39	114	-0.16336	72	189.0	60.3	128.72	213.38%
114	-0.16336	72	123	-0.05027	75	43.9	44.4	-0.53	-1.18%
123	-0.05027	75	134	0.08796	77	51.82	50.8	1.02	2.01%
134	0.08796	77	145	0.22619	76	50.91	49.2	1.69	3.44%
145	0.22619	76	155	0.35186	70	48.95	50.8	-1.85	-3.63%

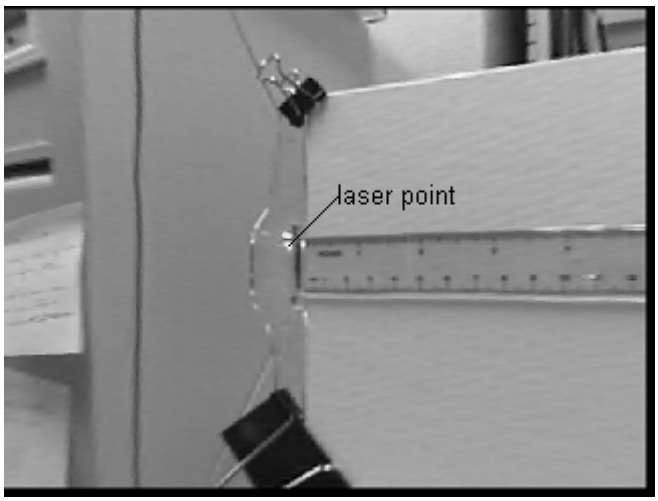

Figure 21: Laser Point - Invalid Reading

The invalid data point is the left-most reading. This reading was taken at the edge of the screen on which the ruler was mounted. The laser point was actually off the edge of the screen. Since the ruler was transparent, this produced an image of the panel behind the screen, so that the laser point seen was shifted to the right (at a distance of 39 pixels from center of image), unlike the remaining readings where the point distance increases and then decreases smoothly. The actual (distortion corrected) image of the first reading

and the two subsequent readings are shown in Figure 21, Figure 22, and Figure 23. This reading shows a potential source of error: the laser point should not move off the object, all readings should be with the point on the object.

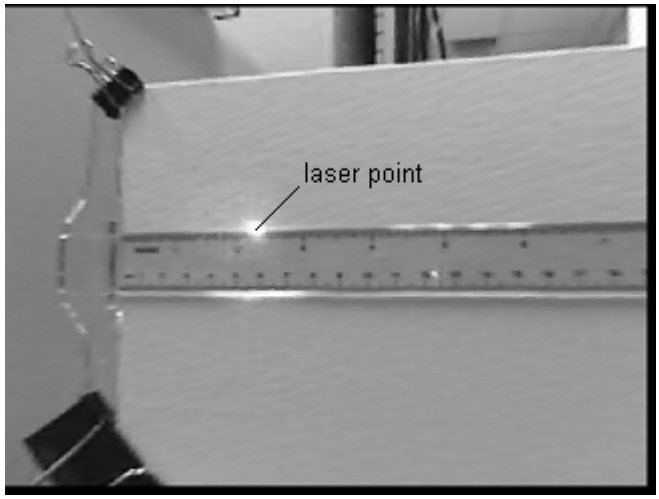

Figure 22: Laser Point Reading 2

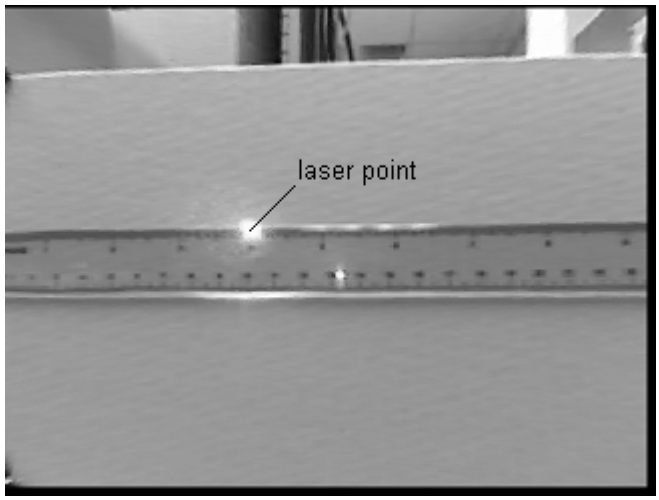

Figure 23: Laser Point Reading 3

Long Range Experiments

The long-range experiments were done to check how the camera scales up to longer distances. We found that while the short range had not required any compensation for alignment errors between camera and laser, the long-range experiments did. The results of the long-range experiments are detailed in TABLE IV below:

TABLE IV

SUMMARY OF LONG RANGE EXPERIMENT RESULTS

Set #	# of meas.	Screen distance (approx) (mm)	Avg Meas. (mm)	Max d (pixels)	Avg. Error (%)	Std. Dev.(%)
1	9	4115	200	5	-17.67	11.72
2	9	3658	200	5	-12.30	8.36
3	9	3962	200	5	-18.03	5.61
4	9	4724	200	5	-28.96	8.46
5	9	4877	200	5	-22.42	10.57

TABLE V

SET 1: LASER MISALIGNMENT UNCORRECTED

Motor angle 1	Angle 1 in radians	Motor angle 2	Angle 2 in radians	Motor angle diff.	Meas. dist. (mm)	Actual dist. (mm)	Error (mm)	Error (%)
14	0.17593	11	0.13823	3	117	200	-83.4	-41.70
11	0.13823	7	0.08796	4	155	200	-44.5	-22.27
7	0.08796	4	0.05027	3	117	200	-83.4	-41.70
4	0.05027	1	0.01257	3	117	200	-83.4	-41.70
1	0.01257	-2	-0.02513	3	117	200	-83.4	-41.70
-2	-0.02513	-6	-0.07540	4	155	200	-44.5	-22.27
-6	-0.07540	-9	-0.11310	3	117	200	-83.4	-41.70
-9	-0.11310	-12	-0.15080	3	117	200	-83.4	-41.70

The results of the long-range experiments show the effects of several of the errors inherent in the system. A slight rotation of one motor unit could change the error from 1% to 24%. This can be seen from the first set of readings, taken at a distance of

approximately 13 feet, with the misalignment error uncorrected and corrected (TABLE V, TABLE VI)

TABLE VI

SET 1: LASER MISALIGNMENT CORRECTED

Motor angle 1	Angle 1 in radians	Motor angle 2	Angle 2 in radians	Motor angle diff.	Meas. dist. (mm)	Actual dist. (mm)	Error (mm)	Error (%)
14	0.17593	11	0.13823	3	152	200	-48.0	-24.0
11	0.13823	7	0.08796	4	203	200	2.6	1.3
7	0.08796	4	0.05026	3	152	200	-48.0	-24.0
4	0.05026	1	0.01257	3	152	200	-48.0	-24.0
1	0.01257	-2	-0.02513	3	152	200	-48.0	-24.0
-2	-0.02513	-6	-0.07540	4	203	200	2.6	1.3
-6	-0.07540	-9	-0.11310	3	152	200	-48.0	-24.0
-9	-0.11310	-12	-0.15080	3	152	200	-48.0	-24.0

Five sources of error could be seen in these results:

1. Laser pointer – camera axis misalignment

   The laser pointer and the camera axis were not perfectly aligned. This introduced an error which had not been very visible in the close range experiments, where the image of the laser pointer would be closer to the image center than it should have been. The computations to overcome this error have been described in the previous section. In these experiments, $\tan\phi$ was determined empirically from calculations on a small subset of the readings. This was because explicit measurement did not provide reliable results since ϕ was small, and the error in the measurements itself was on the same order as ϕ. The result of correcting ϕ on the measurements at 13 feet are shown in TABLE VI.

2. Human element

   Because the actual positioning of the robot is done by a human being, using the intended position for the laser point in the image as a reference, errors in positioning the laser point correctly may occur. These errors are usually because of missing the intended position, or being slightly off the position but not detecting the fact in the image.

3. Finite resolution of the motor

   As the separation of the camera from the object being viewed increases, it becomes increasingly more difficult to rotate the robot to the correct amount to coincide with the edge positions on the object being measured. This error became especially visible in the long-range experiments, as can be seen in the table shown previously (TABLE V). An attempt was made to reduce the problem by halving the resolution of the motor for the long-range experiments(Sets 3-5 used the halved resolution, of 0.36 degrees per motor unit. This also meant the range of the motor was halved, from 180 degrees to 90 degrees). This did not produce a noticeable improvement in the results, as can be seen from the summary in TABLE IV.

4. Motor errors

   Because the motor uses proportional control, the torque applied to move it by a single unit was sometimes insufficient to overcome the static friction. In this case, care had to be taken that movements were done in such a way that the robot was not commanded to move just a single unit when starting from rest. Also, it was found that the robot would not return to the original position the first time it was moved

after being turned on and then given the command to return to the original position. Instead, it would stop at a position a little displaced from the original. This is probably because of static friction in the motor. We also found the motors to have some backlash error. This ranged from one to two units of motion in the opposite direction. Moving in a single direction while taking measurements, and ensuring that a first set of moves were made and the discarded before taking readings helped to minimize both these errors. This strategy was followed while taking all the readings in these experiments.

5. Finite camera resolution

The finite camera resolution causes problems in longer distances, because the rate of change in the angle subtended by the laser spot on the object becomes less as the distance grows. Therefore, when the distance of the camera from the object is 6 feet or further, the distance of the laser point from the center of the image (d) tends to stay the same for increasing distance of separation, until it reduces by a pixel and then stays the same again for some more distance. This problem has the result that the average error within the range of the distance where d remains constant tends to increase from about 12% to about 30% as the distance is increased. One way to counter this error may be by intentionally adding a misalignment to the laser pointer, with the laser pointing outwards rather than inwards as in the case here. Then, as the object went further away, the laser point would move more in the image than it would have for the case of parallel laser pointer-camera axis. Knowing

the angle of the laser pointer with respect to the camera, the distance from the object could be calculated using the modified method shown previously.

DISCUSSION

The accuracy considerations in USAR are not as critical as those in more controlled environments, like industrial manufacturing. Usually, the objective is to simply get a general idea of the size of the object being viewed. The results provided in the previous section show that the method described here achieves this goal. The computation required is minimal, and has few computationally expensive operations. Including the time taken to acquire data, the entire process could be performed in a few seconds per object to be measured. This section describes the modifications needed to use this system for more extensive measurements, and future work and improvements possible on the system.

As stated in the Scale Determination section, with a single laser pointer, there is at least one plane of camera motion along which the distance between the camera and laser would never vary: when the line joining the laser pointer to the camera axis is parallel to the plane of motion of the camera-pointer system. In actual practice, this limitation requires that two laser pointers be mounted on the robot, the second laser pointer being placed perpendicular to the plane of motion of the first camera-pointer system. The output from the computations corresponding to each pointer would then be weighted dependent on the plane of motion of the camera, and combined to produce the end answer to be given to the user.

As mentioned in the long-range experiment results, there were five major sources of error identified. These are examined further here, with possible improvements that

could overcome the errors. In addition, certain errors that were not evident in the experiments described here, but may be present, are also discussed

1. Laser pointer – camera axis misalignment: Imperfect alignment of the laser pointer with respect to the camera axis introduced an error where the image of the laser pointer would be closer to the image center than it should have been. The computations to overcome this error involve finding the error angle ϕ, and factoring this into the computations to find the distance of the object from the robot, as described in the section on experiment setup.
2. Accounting for the human element: Human errors in positioning the laser point correctly often occur. An edge-detection algorithm could be used to find the nearest edge to the laser pointer on each side, and the edge data from the algorithm could be used to provide a closer estimate of the actual size of the object being measured from edge to edge. A mechanism that automatically moves the laser point to the detected edge would improve the process further.
3. Finite motor resolution: Servomotors with a higher resolution are available for a higher price, so this category of error can be reduced by purchasing an appropriate motor.
4. Motor errors due to friction: The robot would not return to the original position the first time it was moved after being turned on and then given the command to return to the original position. Instead, it would stop at a position a little displaced from the original, possibly because of static

friction in the motor. To characterize this error more accurately, some more error measurements need to be made. The first is a measurement of the displacement of the returned-to-origin position has with respect to the original position. the second is a measurement of the repeatability of moves made from the original to a displaced position, every time the robot is turned on. This would help us to know the maximum error that can be expected. In these experiments, this error was minimized by making several movements before beginning the experiment proper. Purchasing a higher quality motor with less friction and/or more gain would reduce this category of error.

5. Finite camera resolution: The finite camera resolution causes problems over longer distances, because the rate of change in the angle subtended by the laser spot on the object becomes less as the distance grows. One way to counter this error is by intentionally adding a misalignment to the laser pointer, with the laser pointing outwards rather than inwards, as was the case here. Then, as the object went further away, the laser point would move more in the image than it would have for the case of a parallel laser pointer-camera axis system. Knowing the angle of the laser pointer with respect to the camera, the distance from the object could be calculated using the modified method shown previously. Additionally, one of the factors that determine the angle subtended is the separation of the camera from the laser pointer. In the setup used for these measurements, this separation was just about 3 cm. This is a very small separation, and could be increased to 10 cm

or more, and still be feasible for mounting on a USAR robot. This would increase the distance for which the measurements remain reliable. A higher resolution camera would also reduce this problem considerably.

6. In addition to these errors, it is possible that the camera axis itself is not parallel to the link axis, and has a misalignment error. To some extent, this error gets hidden by the calibration step performed, since the calibration step uses the camera position to compute the relationship between the camera and link positions. However, at further distances, this error may have an adverse affect on measurements. The experimental setup used here had had limited degrees of freedom, and several errors caused mainly by the hardware limitations of the equipment. These problems masked this error, so that the system was unable to show the error clearly. Therefore, as future work, the effect of this error on the system described needs to be investigated, and it should either be compensated for, or its effect incorporated into the computations.

As a part of the future work, it might be also useful to map the exact distance at which the error deteriorates beyond 10%. This would help in devising strategies to work around the error. Since this is specific to the robot, a robot with better motors and a better camera would see this deterioration much further away, increasing the range for which the robot can be successfully used. An additional direction to investigate further would be the change in the number of pixels between the camera axis and the laser point center as distance from the object increases. This could be mapped by moving the

object further away, and mapping the measurement error with respect to the pixel width at successively greater distances. The variation in the error in the range where the pixel width remains constant should also be investigated.

From the discussion above, it can be seen that the system is limited by the resolution of the camera, as this sets a limit on the smallest difference between subtended angles that can be detected by the camera. This in turn means that as distances increase, the error in the estimate would increase correspondingly, because the camera would be unable to differentiate between increasingly longer distances. The methods described above to reduce this problem can improve the accuracy of the system for longer distances, but cannot eliminate the problem. While increasing the resolution of the camera would help, there is a limit on how much of an increase is feasible, because of the limited communication bandwidth, which would prevent high resolution images being sent back to the robot. However, USAR robots are intended for use in confined spaces within collapsed buildings and other similar structures. This limits the range of unobstructed view the robot would have to a few feet. Therefore, this system can be mounted on USAR robots and should function satisfactorily in that environment.

CONCLUSION

Disasters are unpredictable and can happen any time. When they happen in populated areas, people are often trapped within collapsed buildings. Such people have to be extracted by personnel trained for search and rescue. Since structures still partially standing at a disaster site are often unstable and prone to further collapse, rescue operations pose many hazards to the rescuers.

Robots can be introduced to search and rescue operations to help in performing the reconnaissance, especially of unsafe areas where it would be dangerous for humans to enter. Several challenges need to be overcome before robots can be introduced successfully in this area. One notable challenge is the lack of scale information in the data sent by current robots when on a reconnaissance mission, making it difficult for rescuers viewing the data to recognize objects in the images.

This thesis presents a method for the operator of the robot to get an estimate of the size of the object being seen by the robot. An experimental setup to test this method has been described, and the results have been shown. The results show that a general idea of the size of the object can be given to the operator when the operator moves a laser-pointer camera system on the robot from one end of the object to the other, indicating the two ends to the robot. It should therefore be possible to use this system successfully on one of the present Urban Search and Rescue robots, with few modifications.

Finally, several methods to improve on the work described here have been presented. Such improvements could increase the accuracy of the system beyond what has been experimentally measured here.

REFERENCES

[1] R. Murphy, J. Casper, J. Hyams, M. Micire, and B. Minten, "Mobility and Sensing Demands in USAR," in *IECON 2K: Session on Rescue Engineering*, vol. 1, pp. 138-142, 2000.

[2] H. Kitano, S. Tadokoro, I. Noda, H. Matsubara, T. Takahashi, A. Shinjou, and S. Shimada, "Robocup Rescue: Search and Rescue in Large-Scale Disasters as a Domain for Autonomous Agents Research," in *Proc. IEEE Conf. on Systems, Man, and Cybernetics*, vol. 6, pp. 739-743, 1999.

[3] J. Casper, "Human-Robot Interactions During the Robot-Assisted Urban Search and Rescue Response at the World Trade Center," M.S. thesis, Dept. Comp. Sc. and Eng., Univ. South Florida, Tampa, FL 2002.

[4] R. Murphy, "Gaps in Rescue Robotics," presented at *IEEE Workshop on Safety, Security and Rescue Robotics*, Tampa, FL, 2003.

[5] Y. Aloimonos, *Active Vision*. Hillsdale, NJ: Erlbaum, 1993.

[6] *US&R - National Urban Search and Rescue Response System,* FEMA, Available: http://www.fema.gov/usr/, Accessed: Dec. 2003.

[7] R. R. Murphy, J. Casper, and M. Micire, "Potential Tasks and Research Issues for Mobile Robots in RoboCup Rescue," ***RoboCup****-2000: Robot Soccer World Cup IV*, Lecture Notes in Computer Science, vol. 2019, pp. 339-344, 2001.

[8] S. Stover, "FEMA Rescue Experiences," presented at *IEEE Workshop on Safety, Security and Rescue Robotics*, Tampa, FL, 2003.

[9] G. Kuepper, "Airport Security," presented at *IEEE Workshop on Safety, Security and Rescue Robotics*, Tampa, FL, 2003.

[10] M. Micire, "Analysis of the Robotic-Assisted Search and Rescue Response to the World Trade Center Disaster," M.S. thesis, Dept. of Comp. Sc. and Eng. Univ. South Florida, Tampa, FL, 2002.

[11] R. R. Murphy, "Marsupial and Shape-shifting Robots for Urban Search and Rescue," *IEEE Intelligent Systems*, vol. 15, pp. 14-19, 2000.

[12] R. R. Murphy, M. Ausmus, M. Bugajska, T. Ellis, T. Johnson, N. Kelley, J. Kiefer and L. Pollock, "Marsupial-like Mobile Robot Societies," in *Proc. 3rd Int. Conf. on Autonomous Agents*, pp 364-365, Seattle, WA, 1999.

[13] A. Castano, W.-M. Shen, and P. Will, "CONRO: Towards Miniature Self-Sufficient Metamorphic Robots," *Autonomous Robots*, vol. 8, #3, pp. 309-324, 2000.

[14] S. Hirose, "Snake, Walking and Group Robots for Super Mechano-System," in *Proc. IEEE International Conference on Systems, Man, and Cybernetics*, vol. 3, pp. 129-133, 1999.

[15] A. Wolf, H. B. Brown Jr., R. Casciola, A. Costa, M. Schwerin, A. Shamas, and H. Choset, "A Mobile Hyper Redundant Mechanism for Search and Rescue Tasks," in *Proc. IEEE Intl. Conf. on Intelligent Robots and Systems*, vol. 3, pp. 2889-2895, 2003.

[16] P. E. Rybski, I. Burt, A. Drenner, B. Kratochvil, C. McMillen, S. Stoeter, K. Stubbs, M. Gini, and N. Papanikolopoulos, "Evaluation of the Scout Robot for

Urban Search and Rescue," in *Proc. of AAAI 2001 Mobile Robot Competition and Exhibition Workshop*, pp. 17-22, Seattle, WA, 2001.

[17] "NC State Engineers Design Pipe-Crawling Robot to Save Lives," in *Engineering News*, North Carolina State University, 1999.

[18] H. Amano, K. Osuka, and T. Tran, "Development of Vertically Moving Robot with Gripping Handrails for Fire Fighting," in *Proc. IEEE Int. Conf. on Intelligent Robots and Systems,* vol. 2, pp. 661-667, 2001.

[19] H. Kitano, M. Asada, Y. Kuniyoshi, I. Noda, and E. Osawa, "Robocup: The Robot World Cup Initiative," *Proc. 1st Int. Conf. on Autonomous Agents*, pp. 340-347, 1997.

[20] R. Murphy, J. Blitch, and J. Casper, "RoboCup/AAAI Urban Search and Rescue Events: Reality and Competition," *AI Magazine*, vol. 23 #1, pp. 37-42, 2002.

[21] A. Jacoff, E. Messina, and J. Evans, "A Standard Test Course for Urban Search and Rescue Robots" in *Proc of Performance Metrics for Intelligent Systems Workshop*, Available: http://www.isd.mel.nist.gov/documents/jacoff/USARPerMIS.pdf, 2000.

[22] R. Murphy, J. Casper, M. Micire, and J. Hyams, "Assessment of the NIST Standard Test Bed for Urban Search and Rescue," *NIST Workshop on Performance Metrics for Intelligent Systems*, Gaithersburg, MD, Available: http://crasar.csee.usf.edu/research/Publications/CRASAR-TR2000-12.pdf, 2000.

[23] P. E. Keller, L. J. Kangas, L. H. Liden, S. Hashem, and R. T. Kouzes, "Electronic Noses and Their Applications," in *IEEE Northcon/Technical Applications Conference*, pp. 116-119, Portland, OR, 1995.

[24] *Infrared Camera Systems*, FLIR Systems, Available: http://www.flir.com/, Accessed: Sept. 2003

[25] S. K. Nayar, "Omnidirectional Vision," presented at *International Symposium on Robotics Research*, Japan, 1997.

[26] J. Guehring, "Dense 3-D Surface Acquisition By Structured Light Using Off-the-Shelf Components," in *Proc. Videometrics and Optical Methods for 3D Shape Measurement,* vol. 4309, pp. 220-231, SPIE, 2001.

[27] L. Guisser, R. Payrissat, and S. Castan, "A New 3-D Surface Measurement System Using a Structured Light," in *Proc. IEEE Computer Vision and Pattern Recognition*, pp. 784-786, 1992.

[28] C. Brenner, J. Boehm, and J. Guehring, "Photogrammetric Calibration and Accuracy Evaluation of a Cross-Pattern Stripe Projector," in *SPIE Conference on Videometrics*, vol. 3641, pp. 164-172, 1999.

[29] C. Brenner, J. Boehm, and J. Guehring, "An Experimental Measurement System for Industrial Inspection of 3D Parts," in *SPIE Machine Vision Systems for Inspection and Metrology VII*, vol. 3521, pp. 237-247, 1998.

[30] G. Hu, "3-D Surface Solution Using Structured Light and Constraint Propagation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, #4, pp. 390-402, 1989.

[31] K. Hyun and L. A. Gerhardt, "The Use of Laser Structured Light for 3D Surface Measurement and Inspection," in *Proc. 4th Int. Conf. on Computer Integrated Manufacturing and Automation Technology,* pp. 215-221, 1994.

[32] J. Hyams, M. W. Powell, and R. Murphy, "Cooperative Navigation of Micro-Rovers Using Color Segmentation," in *IEEE International Symposium on Computational Intelligence in Robotics and Automation*, pp. 195-201, 1999.

[33] T. Lindeberg, "Scale-space: A framework for handling image structures at multiple scales," in *Proc. CERN School of Computing*, 1996.

[34] D. G. Lowe, "Three-Dimensional Object Recognition from Single Two-Dimensional Images," *Artificial Intelligence*, vol. 31, pp. 355-395, 1987.

[35] S. F. Ray, *Applied Photographic Optics*, Third ed: Focal Press, Oxford, UK 2002.

[36] R. Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3-D Machine Vision Metrology Using Off-the-Shelf Cameras and Lenses," *IEEE Journal of Robotics and Automation*, vol. RA-3, #4, pp. 323-344, 1987.

[37] *Manual of Photogrammetry*, pp. 801-803, American Society of Photogrammetry, Falls Church, VA, 1980.

[38] R. K. Lenz and R. Y. Tsai, "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, #5, pp. 713-720, 1988.

[39] R. Y. Tsai and R. K. Lenz, "A New Technique for Fully Autonomous and Efficient 3D Robotics Hand/Eye Calibration," *IEEE Transactions on Robotics and Automation*, vol. 5, #3, pp. 345-358, 1989.

[40] T. Thormaehlen, H. Broszio, and I. Wassermann, "Robust Line-Based Calibration of Lens Distortion from a Single View," in *Proc Computer Vision / Computer Graphics Collaboration for Model-based Imaging, Rendering, Image Analysis and Graphical Special Effects (Mirage)*, pp. 105-112, 2003.

[41] D. E. Stevenson and M. M. Fleck, "Robot Aerobics: Four Easy Steps to a More Flexible Calibration," in *Proc IEEE 5th Int. Conf. on Computer Vision*, pp. 34-39, Cambridge, MA, 1995.

[42] G. P. Stein, "Lens Distortion Calibration Using Point Correspondence," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 602-608, 1997.

[43] W. B. Seales and D. Eggert, "Active-Camera Calibration Using Iterative Image Feature Localization," Technical Report, Univ. of Kentucky, Lexington, KY, 1995.

[44] Q.-T. Luong and O. D. Faugeras, "An Optimization Framework for Efficient Self-Calibration and Motion Determination," in *Proc. Int. Conf. on Pattern Recognition (ICPR94)*, pp. 248-252, Jerusalem, Israel, 1994.

[45] O. D. Faugeras, Q.-T. Luong, and S. J. Maybank, "Camera Self-Calibration: Theory and Experiments," in *Proc. 2nd European Conference on Computer Vision*, pp. 321-334, Santa-Margherita, Italy, 1992.

[46] F. Devernay and O. D. Faugeras, "Automatic Calibration and Removal of Distortion from Scenes of Structured Environments," in *SPIE Proceedings - Investigative and Trial Image Processing*, vol. 2567, pp. 62-72, San Diego, CA, 1995.

[47] D. C. Brown, "Close-range Camera Calibration," *Photogrammetric Engineering*, vol. 37, pp. 855-866., 1971.

[48] N. Andreff, R. Horaud, and B. Espiau, "Robot Hand-Eye Calibration Using Structure-from-Motion," *International Journal of Robotics Research*, vol. 20, pp. 228-248, 2001.

[49] *LEXAN(R) polycarbonate resin*, General Electric, Available: http://www.geplastics.com/gelexan/, Accessed: Jan. 2004.

[50] *Polycarbonate Information*, LynxMotion, Available http://www.lynxmotion.com/images/html/infolexa.htm, 2003.

[51] *Serial Servo Controllers*, Scott Edwards Electronics Inc. Available: http://www.seetron.com/ssc.htm, Accessed: Sept. 2003.

[52] *Color CCD Pinhole Video Camera*, CD3 Security, Available: http://www.eyespyvideo.com/color.html, Accessed: Sept. 2003.

[53] I. L. Livshits-Anitropova, M. M. Russinov, and I. G. Bronchtein, "Wide-angle zoom lens with removed forward entrance pupil," in *Proc. Int. Optical Design Conference*, vol. 3482, pp. 233-239, SPIE, 1998.

APPENDIX

This appendix details the work done to remove distortion from the image using the method described in [36]. First, I provide a brief outline of the method as given in the paper, and then I provide the methodology I used to derive the factors I needed. Finally, I provide an analysis of the possible reasons why the method did not work for the experiment setup with which I was working.

Using the nomenclature in Tsai's paper, the process of acquiring an image is as follows:

World coordinates $[x_w, y_w, z_w]$ are transformed and rotated into camera coordinates [x,y,z]. In our case, since the only objective was to correct the distortion, the world and camera coordinates were taken to be coincident.

These coordinates are then combined with the camera focal length to get an ideal, undistorted camera image $[X_u, Y_u]$.

$$X_u = fx/z \tag{65}$$

$$Y_u = fy/z \tag{66}$$

Radial distortion, having coefficients k_1 and k_2, is then applied to the image, getting the distorted image $[X_d, Y_d]$. The distortion is modeled as

$$X_d + D_x = X_u, \tag{66}$$

$$Y_d + D_y = Y_u, \tag{67}$$

where D_x and D_y are the distortion values, given by (62) and (63).

Tsai[36] showed that the first two terms in the distortion equation were sufficient for a radially distorted image to be corrected. The additional terms increased the burden of computation but did not contribute significantly to the results.

The distorted image is then scaled to become the image that forms on the computer. The scale factor is determined by the distance between CCD elements in the camera d_x, d_y in the x direction and y direction respectively. In the case of images being viewed on TV, an additional scale factor S_x was there in the x direction. This was taken as 1 for our computations. Also, the number of pixels in each direction was taken as equal to the number of array elements in each direction.

$$X_f = X_d/d_x + C_x, \tag{68}$$

$$Y_f = Y_d/d_y + C_y, \tag{69}$$

where C_x, C_y are the coordinates of the center of the image.

An image of a calibration grid with points marked at known distances and kept at a known distance from the camera is taken. The above equations are used to work backwards from the image to arrive at the values of f, k_1 and k_2.

From the pairs (62), (66) and (63), (67)

$$X_u = X_d(1+k_1r^2 + k_2r^4), \tag{70}$$

$$Y_u = Y_d(1+k_1r^2 + k_2r^4). \tag{71}$$

From (65), (66), (70), (71)

$$X_d(1 + k_1r^2 + k_2r^4) = fx / z, \tag{72}$$

$$Y_d(1 + k_1r^2 + k_2r^4) = fy / z. \tag{73}$$

Taking two measurements for different values of x and y, with the same z distance, we get two pairs of equations in the form above, one pair for x and one for y. Subtracting the two x equations from each other, we get:

$$(X_{d1} - X_{d2}) + k_1*(X_{d1}* r_1^2 - X_{d2}* r_2^2) + k2*(X_{d1}*r_1^4 - X_{d2}*r_2^4) = (f/z)*(x_1 - x_2).$$

In this equation, we need d_x and d_y to determine X_{d1}, X_{d2}, r_1 and r_2. Since these are not known, we make the assumption that $d_x = d_y = d$. Then, let $X_d/d = X$, we get:

$$(X_1 - X_2) + k_1d^2*(X_1* \acute{r}_1^2 - X_2* \acute{r}_2^2) + k_2d^4*(X_1* \acute{r}_1^4 - X_2* \acute{r}_2^4) = (f/dz)*(x_1 - x_2).$$

X_1 and X_2 are the readings taken from the image obtained. $\acute{r}_1$ and $\acute{r}_2$ are the corresponding r values, given by

$$\acute{r} = (X^2+Y^2)^{1/2}.$$

Now, k_1d^2 and k_2d^4 can each be combined into a single variable k'_1 and k'_2 respectively. This allows the equation to be computed directly from image readings, without modifications. The same operations can be done for the y axis. With four sets of readings, three simultaneous linear equations can be made that can be solved for k_1 and k_2. The readings taken for calculations in x are shown in TABLE VII.

TABLE VII

CALCULATIONS FOR DISTORTION

Set #	Rdng. #	x1	x2	r1	r2	k1	k2
	1	-251	-215	251.128	218.689		
1	2	-215	-171	218.689	193.706	$1.50 * 10^{-6}$	$-4.11 * 10^{-11}$
	3	-171	-123	193.706	229.458		
	1	23	75	25.080	97.309		
2	2	75	122	97.309	162.936	$4.04 * 10^{-6}$	$-3.32 * 10^{-11}$
	3	122	167	162.936	222.481		

Therefore, from the table we have:

Solution 1:

$k'_1 = 1.493989*10^{-06}$

$k'_2 = -4.113665*10^{-11}$

Solution 2:

$k'_1 = 4.042069*10^{-06}$

$k'_2 = -3.320705*10^{-11}$

As can be seen, the results do not agree at all. Therefore, the method was deemed unsuitable for my setup.

There are several possible reasons for this method not having produced usable results. The most likely one is that the camera did not have an ideal form of radial distortion, where the distortion could be modeled directly on the position of the image point with respect to the center of the camera. Another possibility is errors in the location of the calibration points, since the image had been blurred and these points were located by analyzing the image to detect edges and then to detect crosses where the horizontal and vertical edges crossed. A final possibility is that the assumption made that $d_x = d_y = d$ is invalid for the camera I used.

VITA

Maitreyi Nanjanath was born on July 4, 1979 in Palghat, Kerala, India. She received her B.E. in Computer Engineering from Delhi University in the Spring of 2001. She started her Master's degree at Texas A&M University in the Fall of 2001. She also worked as a research assistant for Dr. Volz during her studies at Texas A&M University. Her major fields of interest are Robotics and Artificial Intelligence. She can be reached at

Maitreyi Nanjanath,

E-25 16th Cross Street,

Chennai, Tamil Nadu,

India - 600090