3-D Computer Vision Using Structured Light: Design, Calibration and Implementation Issues[†]

Fred W. DePiero and Mohan M. Trivedi

Computer Vision and Robotics Research Laboratory

Electrical and Computer Engineering Department

The University of Tennessee, Knoxville, TN 37996-2100

depiero@falcon.engr.utk.edu, trivedi@falcon.engr.utk.edu

Abstract

Structured Light (SL) sensing is a well established method of range acquisition for Computer Vision. This chapter provides thorough discussions of design issues, calibration methodologies and implementation schemes for SL sensors. The challenges for SL sensor development are described and a range of approaches are surveyed. A novel SL sensor, **PRIME***, the* **PR***ofile* **I***maging* **M***odul***E** *has recently been developed and is used as a design example in the detailed discussions.*

KEYWORDS: Computer Vision,Range Image Acquisition, Structured Light Ranging, Real-Time Machine Vision, Sensor Calibration

---

[0†]This research is sponsored in part by grants awarded by the Japan Railways and the Office of Technology Development, U.S. Department of Energy.

1 Introduction

Machine vision as a discipline and technology owes its creation, development and growth to digital computers. Without computers machine vision is not possible.

The main objective of machine vision is to extract information useful for performing some task from various forms of image inputs [1, 2]. Examples of machine vision tasks include, robot guidance [3, 4, 5], remote sensing [6], medical diagnosis [7, 8], various types of inspection [9], document processing [10, 11] and many more. For many applications three dimensional (3-D) descriptors of the scene are required. Conventional cameras capture 2-D images and computational approaches are needed to infer the 3-D descriptors from one or more images. Common approaches for this include the use of 2 or more cameras in binocular and photometric stero [12, 13, 14, 15, 16, 17, 18]. Approaches using single images include various structure from "x" (x=shading, texture, shadows,motion) techniques [19]. Recent advances include structure from image streams [20, 21] and frameworks for integrating more than one technique for 3-D information extraction [22].

The above approaches share one very important common feature. They are all "passive" approaches, i.e. they do not need a special source of energy to illuminate the scene. There are obvious advantages of this approach: cost, simplicity of imaging hardware, compatibility with human visual processes, etc. On the other hand these approaches need to also overcome some inherent challenges. These challenges arise from the loss of information associated with the perspective mapping of a 3-D scene onto a 2-D image. This produces a fundamentally ill-posed problem when single images are used to find the 3-D descriptors.

Additional challenges that face machine vision researchers are due to spatial and intensity

quantization of images and due to computational requirements. These effects manifest themselves in robustness, speed and accuracy performance metrics. In some sense, the advantages of simpler and low cost acquisition hardware is compensated with the need for sophisticated computational processing and analysis approaches.

"Active" approaches for 3-D vision use specialized illumination sources and detectors. These techniques overcome the fundamental ambiguities associated with passive approaches. Some active techniques use laser radar [23, 24], others use various forms of structured lighting [25, 26, 27, 28, 29, 30]. In general these methods are able to eliminate the ill-posed problems associated with passive approaches by modeling and observing the projected illumination. In the case of laser radar, a spot laser beam is steered across a scene. Range measurements are made by either AM or FM detection schemes. Structured Light techniques model the optical paths associated with emission and detection to compute range data by triangulation.

Structured Light ranging has some particularly attractive features compared to laser radar approaches. SL systems can be designed "from the ground, up" much more so than with a packaged laser radar sensor. This provides significant advantages for customizing the acquisition capabilities of a sensor for a particular application. Generally speaking, SL sensors are also more accurate and can be made more rugged, and less expensively than laser radar devices [31].

This tutorial focuses on Structured Light sensing. Some critical design issues are discussed, as well as methods of sensor calibration and metrics for calibration models. The implementation of **PRIME**, the **PR**ofile **I**maging **M**odul**E**, is described in detail, including various design tradeoffs and sensor performance benchmarks.

2 Introduction to Structured Light Ranging

Structured Light (SL) sensing is a well established technique for ranging. A great variety of techniques in this general area have been developed [32, 26, 24]. The common thread of all these approaches is the underlying use of triangulation. This ranging geometry can be seen in Figure 1, which depicts the optical components in PRIME. Here, the laser is projected downward towards objects in the scene. Laser illumination striking an object is observed by the camera to reveal surface profiles which can be converted into Cartesian range data. In PRIME, the ranging triangle is completed by a rigid backbone between the camera and laser.

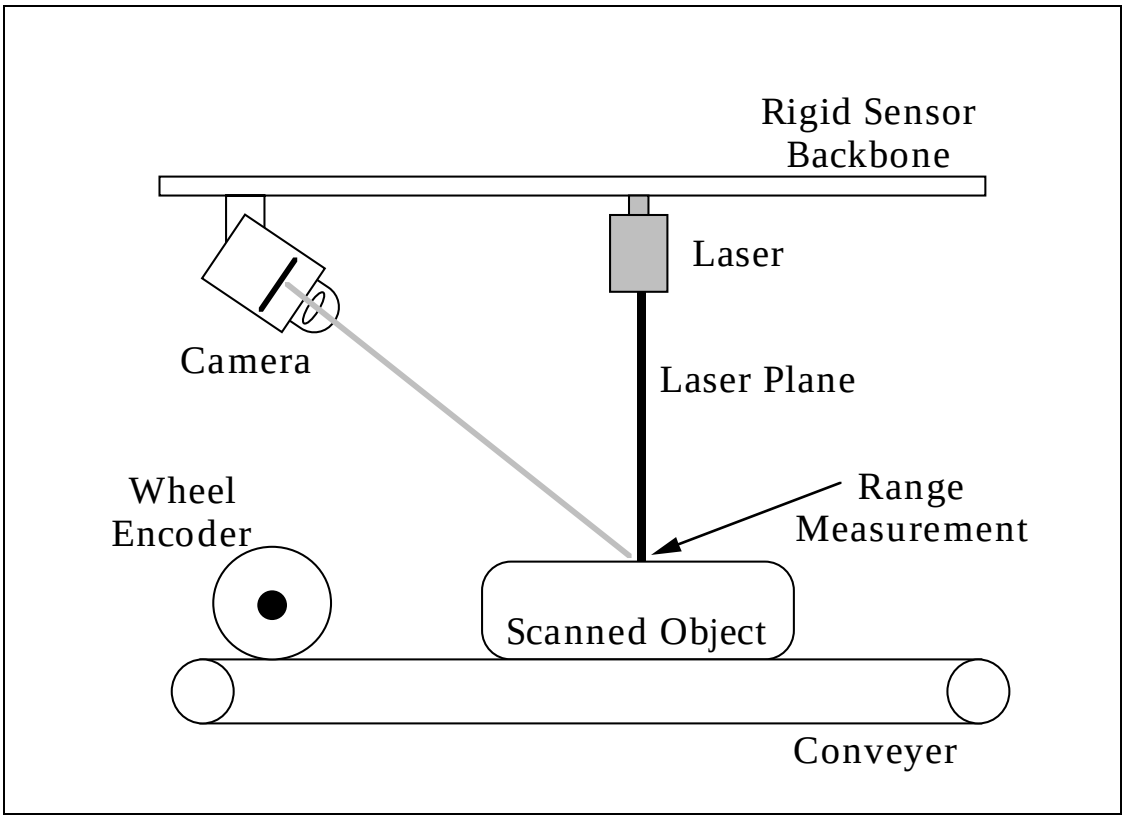

Figure 1: *Structured Light systems use triangulation to acquire range measurements. In PRIME, the ranging geometry is formed by a laser emission, the reflected light observed by the camera and by a rigid backbone. Optical measurements capture range data in a plane. A conveyer produces the necessary motion for 3-D range data.*

The three main challenges in developing a Structured Light system are calibration, accuracy and acquisition speed. Calibration models are needed to relate image to world coordinates. See Figure 2. Methods of calibration are critical for success, but are usually not described in sufficient detail in published works. Typically, procedures require examples of image coordinates to be generated by calibration targets, the positions of which must be known with high accuracy. To ensure accurate ranging in the final system, the calibration data must be generated at distances from targets that span the range of standoffs that will be used during sensing. If the intended standoff is on the order of a few feet or less, then standard optical bench equipment can be of great utility during calibration. Hence, the difficulty of this task can vary with the intended standoff of the sensor.

Challenges in calibration also arise from the complexity of models that are required for sensor kinematics. These kinematics describe the geometrical relationship between the camera and the laser plane. If, for example, these kinematics are fixed then the calibration procedure is much simpler. PRIME is such a system, which is referred to here as having a "fixed-plane geometry", as seen in Figure 1. An alternative to having a fixed triangular geometry is to reorient the optical paths while ranging. Using a "dynamic-plane geometry" in this manner [29, 33] can permit larger regions to be scanned more rapidly. These approaches typically reorient the laser beam using low inertia optical components. Beam repositioning in this manner can be achieved with high speed and high precision. While dynamic geometries are attractive from the point of view of acquisition speed, they usually require more complex calibration models [34, 29, 30].

Ranging accuracy is highly dependent on calibration models. However, even the most careful calibration effort can be fruitless if the ranging geometry is unfavorable. This refers to the

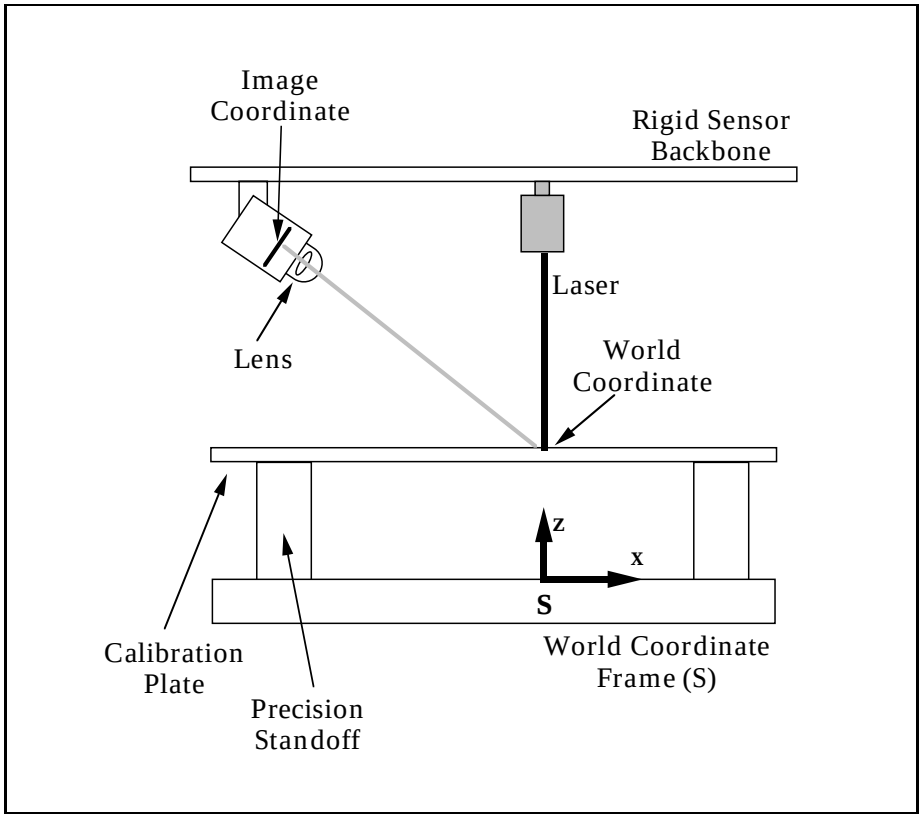

Figure 2: *The purpose of sensor calibration is to find a mapping between image and world coordinates. In PRIME, this relationship is fixed because of the rigid backbone between the laser and camera.*

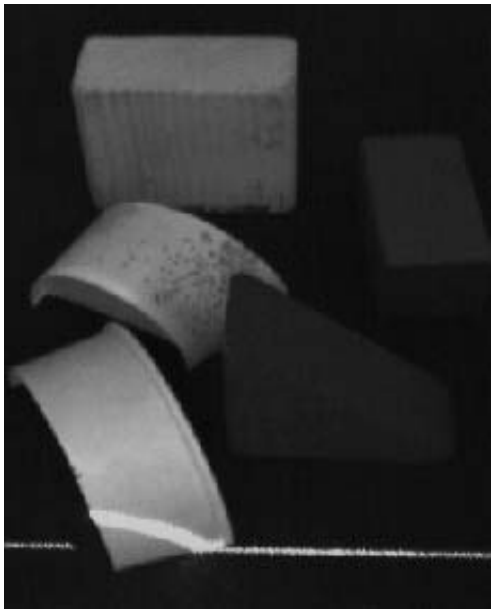

Figure 3: *Image of laser profile seen by camera. The camera's optical filter was removed for this image to better reveal objects in the scene.*

sensitivity of range measurements to various system parameters, see section 6 for a sensitivity analysis of PRIME.

Acquisition speed is another challenge to SL system design. Range acquisition involves locating the laser profile within camera images. This requires pixel examination and processing in order to precisely locate image coordinates that reside at the center of the laser profile. While the complexity of these algorithms are low, the computational requirements are relatively high. As seen in Figure 3, systems such as PRIME must process images with relatively low information density - this image has only one view of the laser line. Given the size of a standard image and relatively small area of pixels illuminated by the laser line ($\approx 1\%$), a significant amount of pixel-level computational effort must be expended to acquire the range data associated with a single image. Because of this relatively low payoff, many researchers have experimented with ways to pack more laser lines into a single image [35, 28, 29]. Approaches with multiple laser images necessitate using a heuristic to determine the correspondence between image features and laser positions. Several methods for this are reviewed below.

Some discussion is appropriate concerning the fundamental limitations of SL ranging. Surface reflectivity is one such factor. For reliable range data, a scanned object should have surfaces with lambertian reflectivity. Specular surfaces will often reflect too much of the structured illumination away from the camera. This produces voids in range data. Note that the degree of surface reflectivity can be counter-intuitive when dealing with near-InfaRed (IR) laser systems, such as PRIME, since these wavelengths are beyond the human visual range.

Shadowing is also a fundamental problem in SL systems. This occurs when object geometries occlude the laser from the field of view of the camera. Shadowing effects can be reduced when

the camera to laser baseline distance is shortened. However, this also increases the sensitivity of the system to measurement noise - see section 6.

Ambient lighting is an important design issue. It can be a limitation to SL or other types of optical ranging if ambient sources are unfavorable and cannot be controlled. Monochromatic illumination and matched optical filters for cameras can be used to tackle this problem, as in PRIME, provided the ambient lighting can be setup on a different wavelength. Florescent lighting and near-IR SL ranging make a very nice complementary pair. This combination provides ample room light without contaminating range imagery. Incandescent light is a very poor choice for use with near-IR SL systems.

The novelty of PRIME stems from new approaches to calibration and to pixel-level operations. These techniques make for a system that has real-time acquisition, is accurate, easy to calibrate and made entirely from commercial components. The following sections review other SL efforts and then detail the approach taken in PRIME.

3 Literature Review and Highlight of Critical Design Issues

The scope of this review is limited to ranging systems with the same style of laser emission as PRIME - a laser plane. Excellent reviews are available for a much broader scope in [24, 23]. The purpose of this review is to highlight design alternatives in Structured Light sensors and to examine the tradeoffs taken in PRIME.

Three critical aspects SL sensor design are presented here. **Backbone Geometry** describes

the geometrical relationship between the laser(s) and camera. **Correspondence** and **Pixel-Level Analysis** are image processing issues at the large and small scale, respectively. The correspondence problem deals with establishing an association between a region of an image and the location of the illumination source (in world coordinates). Pixel-level analysis takes over once the correspondence problem has been solved, to find the precise image location of illumination patterns.

3.1 Backbone Geometry Designs

This aspect of a SL sensor design has to do with the geometry between the camera and laser plane. "Fixed-Plane" geometries have advantages in calibration, simplicity and ruggedness due to the absence of any moving optical components. This is the approach taken in PRIME. This necessitates some other source of motion in order to acquire range data in 3-D. In [36] a fixed-plane is used in conjunction with a rotating table. One of the first industrial applications of SL [32] used a conveyer belt, as does PRIME.

"Dynamic-Plane" systems alter backbone geometry during the ranging process. These systems are generally able to scan larger areas more quickly, and do not require additional mechanisms to supply motion. Accuracy and calibration can become more challenging, however. Rotating optical components began appearing in the work of [37, 38, 39]. In [35] the imprecision associated with using gear trains to rotate optics is described. Other approaches have used direct-drive motors, some galvanometer-based, to rotate optics [34, 29].

The complexity of modeling the laser reflection can vary significantly, depending on the required accuracy of a system. Simple reflection models [33, 40] assume a perfect alignment between

the laser line and the axis of rotation. When a front-silvered mirror is rotated about an axis that is offset from the reflecting surface, a displacement of the laser plane is introduced. This displacement is also sometimes ignored. In [34, 29] a more complete geometrical model is used.

In [41, 30] a half-sphere field of view necessitated the use of a dynamic-plane system with pan and tilt units for both cameras and lasers. Emitter and detector units were also housed separately, requiring an in situ calibration as well as dynamic geometrical models for each unit.

3.2 Approaches to SL Correspondence Problem

It is reasonable to consider imaging multiple laser profiles in a single camera image in order to achieve increased acquisition rates. More than one laser profile per image increases the information content, but necessitates a scheme for establishing the correspondence between a laser profile and the associated laser plane geometry. Having no ancillary means to establish correspondence [42] can result in a combinatorically unfavorable problem. In [25] the correspondence problem is addressed using relaxation labeling.

In [35] a color camera was used to observe multicolored profiles. In this approach the position of each colored plane was fixed and all planes were projected simultaneously. This provided images with a clear correspondence relationship that could be acquired in a single frame time.

In [29] a time-lapse image of closely-spaced profiles was collected. Here the correspondence problem was solved by computing a second registered image that contained position-stamp information for each profile. This was accomplished using simple processing steps, allowing position-stamping to be implemented in real-time on commercially available hardware.

Another useful technique involves collecting a sequence of well registered images while illumi-

nation sources are toggled on and off. In this way, each laser profile can be identified by analyzing which images in the sequence it appears [43, 44, 45, 28].

If multiple laser profiles are present in an image, or if a sequence of images must be analyzed, then it can be necessary for objects in a scene to remain motionless during the acquisition cycle. These types of approaches introduce limitations in more dynamic environments.

3.3 Methods of Pixel-Level Analysis

Once the correspondence between a profile and the actual laser position are established, it is necessary to precisely determine image coordinates at the center of the laser profile. The precision with which these coordinates are located effects the overall accuracy of range measurements. Speed and accuracy tradeoffs exist here.

One approach to speeding up SL acquisition is to perform the "pixel-level" analysis in the analog domain. An early effort in this area [46] made the assumption that stripes are roughly vertical. This approach used dedicated timing hardware to find the illumination on each horizontal scan line. This provided image coordinates at frame rate image. More recent approaches [40] have used VLSI implementations that incorporate analog detection and timing operations in a single chip.

In [35] a peak detection algorithm was used to find image coordinates at the center of the laser profiles. This analysis did not provide results with subpixel accuracy. An analog implementation of this scheme was also proposed.

In [41] the nominal orientation of the laser profile could not be assumed and pixel-level operations had to be performed at different orientations. Here, video images were digitized and

pixel-level analysis was performed in different directions, using an adaptive scheme.

3.4 Alternate Structures for Illumination

It is worth considering a variety of structures for illumination when at the early design stage of a SL system. One such approach uses a spot of laser illumination. These are sometimes referred to as "flying dot" systems.

If a video camera is used for this type of system, the approach can suffer in terms of acquisition speed due to the low information density per image [47]. Because of this, many approaches turn to 1-D detectors [48] or to custom optics [49].

Despite the problem of low information density, there is a significant advantage to flying-dot systems that use 2-D video images, or some other type of 2-D detector. When a directional vector (in R^3) is available, from the detector to the flying dot, an on-line confidence measure can be computed for each range point. The ranging process in these systems can be formulated as an intersection calculation of two lines, one along the laser optical axis and the other along the camera sighting of the flying dot. The closest point of intersection between these two lines can be used as a best estimate for a range measurement. The minimum distance between these two lines can then be used as an estimate of the measurement uncertainty. This confidence measure is provided in an on-line, point-by-point manner. This provides great advantages for applications demanding high accuracy and high reliability. SL systems that image a laser line, as with PRIME, do not possess this type of inherent accuracy check.

Laser optics are available that project alternate light patterns, circular projections, for example. These may provide advantages for some situations where the structure of the light matches

an application-dependent measurement region.

3.5 The PRIME Niche

PRIME has been designed for high accuracy, ruggedness and simple calibration. To accomplish these goals in the most reliable manner and at the highest possible speed, it was decided to use a single plane of laser light. This provided ruggedness, accuracy and simple calibration, but did require effort to achieve real-time acquisition. Acquisition speeds were improved by mapping portions of the pixel-level operations onto dedicated commercial hardware. Because a single laser profile is imaged, and because range acquisition occurs at frame rate, PRIME is able to scan objects that are continuously moving.

4 Structured Light Acquisition

The architecture of a Structured Light sensor includes both optical components and pipelined processing elements. A wide variety of components can be selected for these purposes. In PRIME, for example, the computing components include a Motorola 68040-based single board computer and a Datacube MV20 image processing board. The main processor runs under a vxWorks environment. It is responsible for real time configuration and control of the Datacube hardware, and for applying calibration models.

4.1 Optics and Imaging

Figure 4: *Laser (left) and camera (right) used in PRIME sensor. The laser generates a plane of light using a cylindrical lens. A bandpass optical filter on the camera yields distinct imagery of the laser plane as it intersects objects in the scene.*

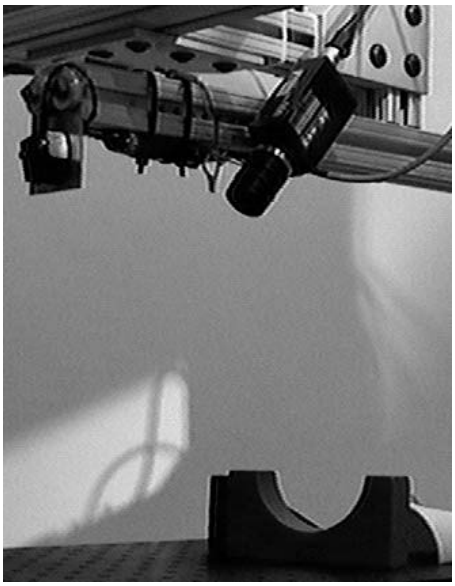

Figure 5: *PRIME sensor with typical objects.*

Figure 4 shows the optical components that comprise the SL sensor. A near-InfaRed diode laser, seen on the left, illuminates a scene from above. The laser emission is in the form of a plane of light which is generated by a cylindrical lens mounted in the laser housing. The black and white camera, seen to the right, is positioned so as to image the light reflected from the laser plane as it strikes objects in the scene. Figure 5 includes typical scanned objects, also.

The camera is outfitted with an optical filter that is matched to the laser optical frequency. These matched optics produce very distinct imagery of the laser profile. The image in Figure 3 was taken with the optical filter removed, for presentation purposes. Figure 10 shows an image captured with the IR bandpass filter installed, as is typical during ranging. Because CCD cameras are quite sensitive to near-IR, manufacturers typically install an IR cut filter. Such a filter was removed from the camera used in PRIME.

Another aspect of the imaging process has to do with the use of an electronic shutter. In applications with objects that move continuously past the sensor, a certain degree of blurring would nominally occur in each camera image. Blurring increases the uncertainty with which image coordinates at the center of the laser profile can be recovered, and hence, must be limited. The camera's electronic shutter reduces the temporal integration period for each pixel. In this way the frame rate of images is unchanged, but the exposure time is reduced. This yields video streams which are subsampled in time, this means that shape information will be missing between sequential laser profiles. This introduces an upper limit on the spatial frequency content of shape descriptions and can result in aliasing.

A standard video frame is composed of two interlaced fields. Each field is transmitted sequentially. When using an electronic shutter, each field is exposed individually. Typically the shutter

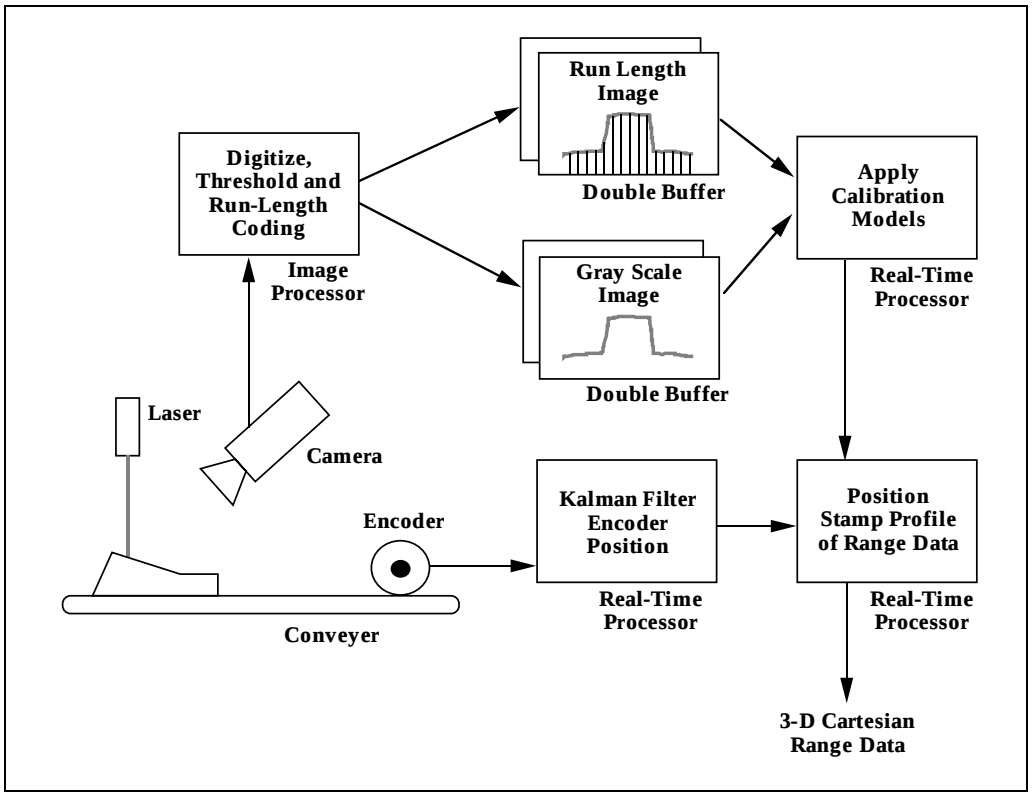

Figure 6: *Block diagram of the processing steps in the PRIME sensor.*

interval for one field will occur immediately before that field is due to be output on the composite video signal. Hence, a single camera image contains data acquired from two separate time instances, 1/60 second apart. This necessitates that the temporally-skewed fields be processed separately when the laser profiles in each are converted to range data.

In PRIME, the physical thickness of the plane of laser light is ≈ 0.040 inches thick. The sensor has been designed for scanning velocities of ≈ 1 in/sec. An electronic shutter interval of $< 1/250$ sec is well suited for this situation because the blurring of the beam is limited to $< 10\%$ of the thickness of the laser plane. In practice this fractional increase in the apparent thickness of the beam has proved to have little effect on range measurements. However, it is true that unfortunate scene geometries can generate greater degrees of blurring. For example, A sloped block will tend to cause the laser line to blur vertically if it is viewed while approaching the camera. Steeper block faces will tend to produce greater degrees of blurring. These factors are very much scene- and application-dependent.

A fundamental tradeoff exists between the ability to localize the position of the laser in an image and the apparent brightness of the beam. The mechanism for adjusting this degree of freedom is the electronic shutter interval. The limiting factor in the range of this adjustment is the amount of optical power in the laser and the sensitivity of the camera. In PRIME, the Automatic Gain Control (AGC) on the camera has been engaged. Since the majority of the camera image is black, the AGC sets analog gains relatively high. This causes the laser illumination to be amplified significantly. For most of the experiments performed with PRIME, the intensity of pixels at the center of the laser line were saturated (256/256).

4.2 Image Processing Operations

The images used during range acquisition have an almost totally black background, with gray levels $\approx$ 10-20 out of 256. Given the saturated pixels on the laser profile, this represents a SNR $\approx$ 25 db for the image data. These conditions provide considerable latitude in selecting a gray scale threshold used for producing binary images. These binary images are used in an intermediate step to roughly locate the laser line within the image.

Image processing operations were split between the Datacube board and the main processor for improved pipeline throughput. In order to permit both devices to simultaneously process images, double buffering was used between these two pipeline elements. See Figure 6. The Datacube board provided image I/O and storage, displayed the real-time status of the ranging sensor, and was used to locate the rough position of the laser line within each image.

The rough position estimates were achieved by thresholding and then run length coding [50]. Run length coding (RLC) generates a 1-D array having indecies with a one-to-one correspondence with the columns of the binary image. Each RLC array element contains the height along a column from the bottom of an image up to the first illuminated pixel. The RLC data is depicted in the image buffers of Figure 6. Hence the RLC array provides a succinct description of the rough location of the laser line. This information greatly improves the speed of range computations because the Datacube board can provide the RLC data at frame rate. This eliminates an otherwise burdensome effort by the main processor of searching the entire image for a relatively small illuminated region - which could not be done in real-time. Using the run length coded array, the main processor can directly access the required portions of images.

To precisely localize the center of the laser line, pixel values are examined on a cross section

of the laser illumination and a weighted centroid calculation is performed. Figure 7 illustrates typical pixel intensities, as sampled along a path roughly orthogonal to the laser line. In PRIME, the intensity profiles are typically 5-7 pixels wide with a roughly Gaussian shape. The main processor examines gray scale pixel values in a 20x1 window, W, centered at locations given by the RLC array. A mapping function, $P(g_i)$, is used to describe the likelihood that a pixel having intensity g_i is a member of the laser line. The mean, $\bar{r}$, and variance, σ_r^2, of the row at the center of the laser line are found with

$$\bar{r} = \frac{\sum P(g_i) r_i}{S}, \qquad \sigma_r^2 = \frac{\sum P(g_i) r_i^2}{S} - \bar{r}^2 \tag{1}$$

where g_i is the gray level of the pixel in row r_i, and $S = \sum P(g_i)$. Summations are taken within the window, W, and all include pixels above the binary threshold. The image coordinate recovered,$(\bar{r}, c)$, is the mean row together with the column under examination.

It was desired to make the mapping function, $P(g_i)$, a smooth curve varying from 0.0 to 1.0, corresponding to pixels in the background and on the laser profile, respectively. A truncated error function (erf) was chosen for the mapping. Values for $P(g_i)$ were found by integrating a Gaussian and then normalizing so that the area under the erf curve was unity. The shape of the original Gaussian was chosen so that 3σ below the mean were gray levels clearly belonging to the background and 3σ above the mean were levels at the center of the laser profile. The gray levels at these 3σ points were chosen manually by examining a histogram of typical images. Previous experiments in camera calibration [34] have shown as much as a 13% change in centroid location when weighted centroid calculations are done, versus binary methods.

The variance, σ_r^2, of each $(\bar{r}, c)$ coordinate was checked against a threshold as a means to eliminate the use of blurry portions of a laser profile. Blurry imagery can result from excess relative motion or uneven reflectivity, for example. In these situations the reliability with which the center of the laser profile can be recovered is compromised and the generation of range points should be avoided in order to maintain accurate results.

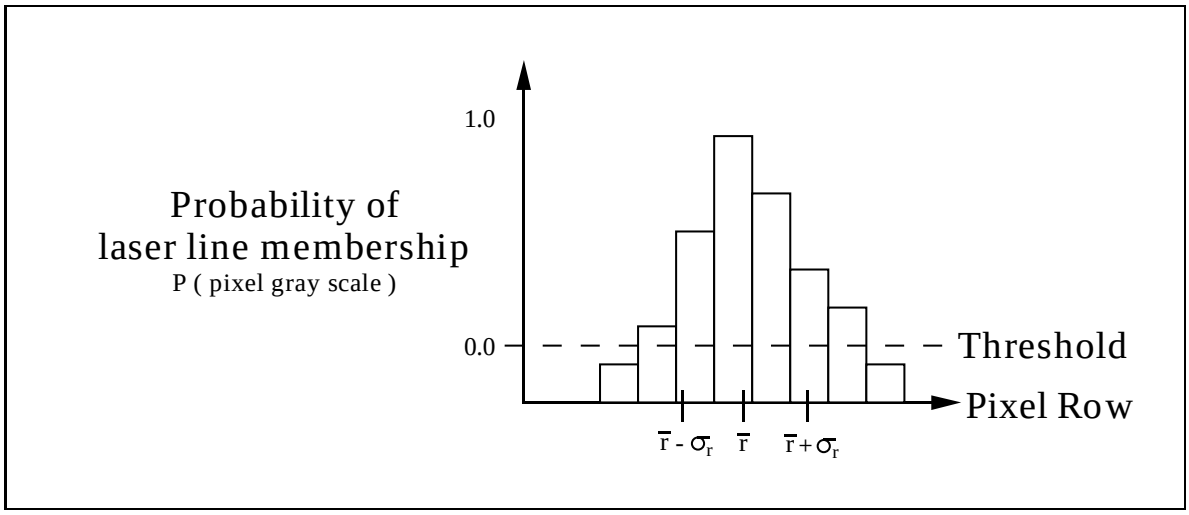

Figure 7: *Recovering image coordinates at the center of the laser line.*

4.3 Ranging Computations

A calibration model, M, is used to convert image coordinates, (r, c), to world coordinates, (y, z). The model is applied to an augmented vector of image coordinates. For example, with a second order model

$$\begin{bmatrix} r^2 & c^2 & rc & r & c & 1 \end{bmatrix} M = z \tag{2}$$

where z is the Z coordinate of a range point. A similar model is used to generate Y coordinates, see section 5.

In PRIME, image data provides the Y and Z components of range points. The X component is found by monitoring the motion of a conveyer belt using a wheel encoder. See Figure 1. The processing of encoder readings is interrupt driven, occurring at a rate four times higher than the video field rate (60 Hz). This provides accurate position estimates (in X) for each video field, allowing the range profiles in each to be field to be properly position-stamped. The main processor was responsible for orchestrating the image acquisition process and stamping each acquired profile with the appropriate positions along the conveyer. Having camera image data synchronized with conveyer position stamps allows 3-D Cartesian range data to be computed. The encoder on the PRIME testbed has a resolution of 4000 counts per revolution, corresponding to 0.003 inches of travel per encoder tick.

The main processor uses Kalman filtering [51] to process the raw encoder readings, for improved position estimates. Because of the relatively constant velocity of the PRIME conveyer, the state transitions of the Kalman filter are modeled as having a constant acceleration [52]. The process noise matrix has been setup accordingly [53]. The process noise parameter has been determined experimentally, $q = 50$. The measurement uncertainty σ_m has been set using the encoder spacing and the standard deviation associated with a uniform distribution [54], $\sigma_m = 0.003/\sqrt{12}$ inches.

Because of the sensor geometry, a simple X-Y grid was used to store range data. This provided advantages in terms of complexity and the speed of data storage. It eliminated the need for an octree data structure, for example. The laser plane has a near-vertical orientation. Because of this, multiple range points do not tend to occur that have the same lateral (X-Y) position above the grid.

5 Calibration of Structured Light Sensors

Calibration models are required in order to provide the relationship between image and world coordinates (see figure 2). A distinction is made here between the problems of calibration versus registration of the sensor. Herein, "calibration" refers to a process in which the location of the world frame (S in the figure) is defined locally to the sensor. Its position is established with the convenience of sensor calibration in mind. "Registration" refers to the process of relating the sensor frame, S, to some other frame that is pertinent to the application - such as a manipulator frame.

In most cases, the design of two aspects of calibration are tightly coupled. These are (1) the process by which calibration data is collected and (2) the formulation of the calibration model. The solution to these two problems typically must be found in a joint manner. The approach taken for PRIME is described below.

5.1 Calibration Process and Formulation of Calibration Model

Structured Light ranging is fundamentally a process of triangulation. Calibration is sometimes approached as a process of isolating explicit geometrical parameters of this ranging triangle. In [24] range calculations are described using the law of sines together with a pin hole model of the camera. Note that this would necessitate two separate calibration procedures (each of which would contribute errors).

A one step calibration procedure has been developed for PRIME. This process is very similar to the Two Planes method of camera calibration [55]. In general, any Fixed-Plane SL system

can be calibrated in a one step procedure because of the rigid mapping between image and world coordinates. One step procedures have advantages in terms of accuracy and simplicity. Accuracy is improved because models can be found via a single least norm solution.

A model $z = f(r, c)$ relating height, z, to image coordinates has been found using empirical calibration data. This relationship has been determined by analyzing images to find many examples of the triplet

$$\left| \begin{array}{ccc} z_i & r_i & c_i \end{array} \right|. \tag{3}$$

Figures 8 and 9 illustrate the PRIME calibration jig. The jig allows a series of horizontal plates to be located at known heights, z_i, and imaged by the camera. Figure 10 shows such an image containing many examples of where the laser plane is imaged at a given vertical height. The pixel-level operations described in section 4.2 are used to generate individual triplets. A number of calibration plates were located across the depth of field of the sensor. Precision standoffs were used to locate these in the Z direction. When calibrating for X, a linear table was used to automate positioning.

For increased image fidelity during calibration, a number of images of a given calibration plate were averaged. This reduces the random noise in pixel values that nominally accompany the imaging process [50].

To find the calibration model, an overdetermined set of equations is formed by augmenting the image coordinates (r_i, c_i) of each triplet. A variety of forms of calibration models have been

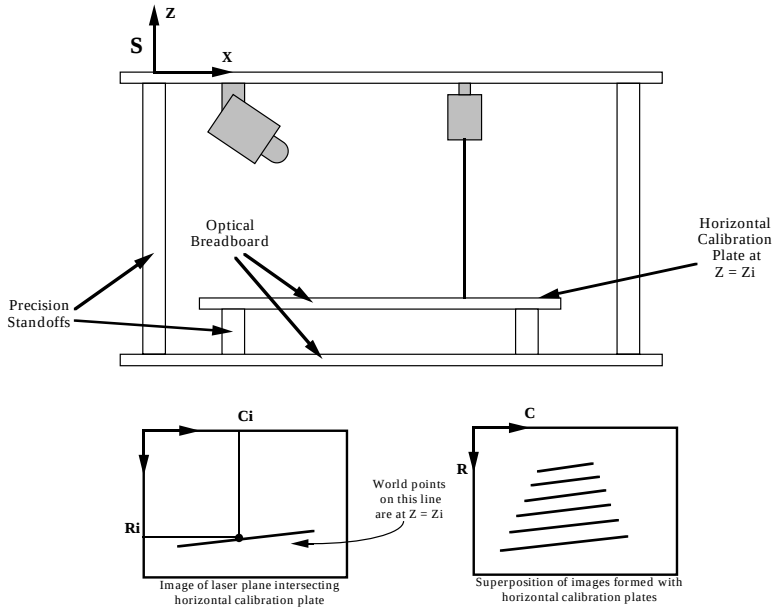

Figure 8: *A series of horizontal plates are used to generate examples of the* $\left| \begin{array}{ccc} z_i & r_i & c_i \end{array} \right|$ *triplet (Eq. 3) for PRIME. The triplets are used to form an overdetermined set of equations from which calibration models may be determined. This illustrates the calibration procedure for the* $z = f(r, c)$ *model.*

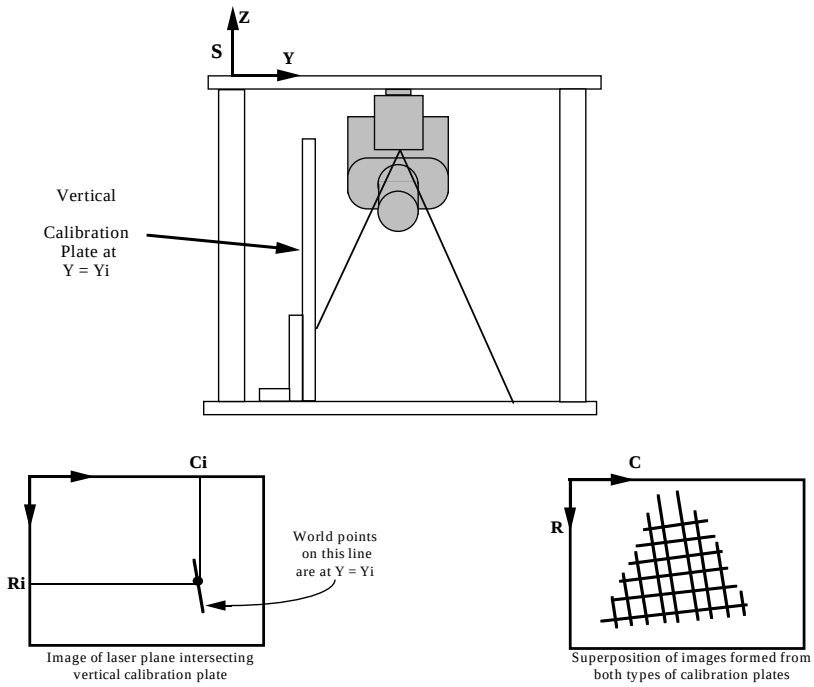

Figure 9: *A series of vertical plates are used to generate the data needed for the $y = f(r, c)$ calibration model in PRIME.*

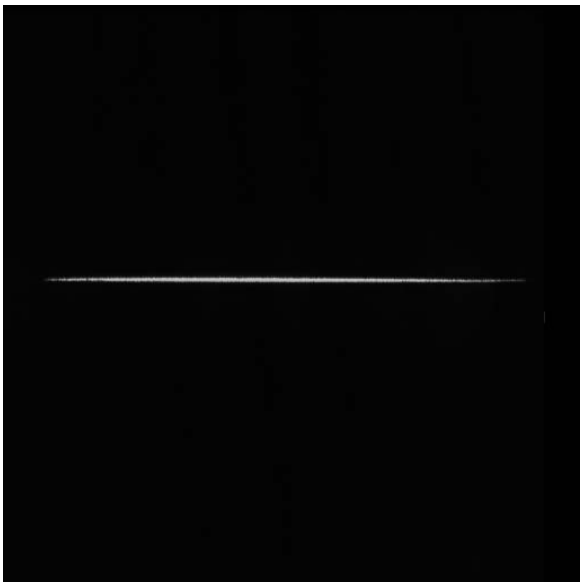

Figure 10: *An image used for range measurement calibration. The central horizontal line is produced by the intersection of the laser plane with a horizontal metal plate.*

studied. In the case of a second order model, for example, a calibration n-tuple

$$\begin{vmatrix} z_i & r_i^2 & c_i^2 & r_i c_i & r_i & c_i & 1 \end{vmatrix}. \tag{4}$$

is used. These n-tuples are arranged to form

$$\begin{vmatrix} z_1 \\ . \\ z_i \\ . \\ z_n \end{vmatrix} = \begin{vmatrix} r_1^2 & c_1^2 & r_1 c_1 & r_1 & c_1 & 1 \\ \cdots & & & & & \\ r_i^2 & c_i^2 & r_i c_i & r_i & c_i & 1 \\ \cdots & & & & & \\ r_n^2 & c_n^2 & r_n c_n & r_n & c_n & 1 \end{vmatrix} \begin{vmatrix} m_1 \\ \cdots \\ m_p \end{vmatrix} \tag{5}$$

or

$$Z = FM \tag{6}$$

Where

Z (nx1) :contains world coordinates,

F (nxp) :contains augmented pixel coordinates, and

M (px1) :contains model parameters.

In this case $p = 6$. A series of vertical plates are used to generate the calibration triplets needed for the Y model. This process is depicted in Figure 9.

5.2 Numerical Method of Solution for Calibration Models

Many methods of solution are possible for finding the calibration model, M, in a least norm sense [56]. However, prudent selection of a solution technique can have side benefits in the calibration process. Forming the normal equations [57], for example, allow the solution to be found, but do not provide any additional information. Techniques that reveal the numerical sensitivity of the solution are much more desirable. This provides an indication of how much the solution will change due to slight perturbations in the calibration triplets. The calibration triplets will contain measurement noise, hence the need for an overdetermined specification of the solution and for as low a numerical sensitivity as possible. There is no guarantee that the existence of ill-conditioned calibration data will be apparent in the residual error [58] of the solution. The greatest care in the imaging and pixel-level operations will be fruitless if the solution for a calibration model suffers numerically. The consequence for ranging is that a calibration model may not yield accurate results when new data is applied.

A common reason for the matrix F of Eq. 6 to become ill-conditioned occurs due to a poor distribution of the input data points. As seen in Eq. 5, the right hand column of entries in F are all 1. The columns of F must be linearly independent [57]. Hence if all calibration triplets possess the same row or column then F will become ill-conditioned.

One technique for improving numerical performance when finding M involves scaling the image coordinates of the calibration triplets. By normalizing the image coordinates to the range $(0, 1)$, the span of entries in F can be reduced. This improves the condition of F [56]. If it is assumed that coordinates vary from 5 to 90% of the image then the normalized and unnormalized coordinates cover the ranges $(0.05, 0.95)$ and $(25, 485)$, respectively. The span of largest to smallest

Table 1: *Demands on numerical precision for normalized and unnormalized calibration models.*

Order	Span of Entries in F Unnormalized	Span of Entries in F Normalized	Cost Ratio of Spans	Integer Precision of Cost (bits)
1st	$485^1/1$	$1/0.05^1$	$2.4 \quad 10^1$	5
2nd	$485^2/1$	$1/0.05^2$	$5.8 \quad 10^2$	10
3rd	$485^3/1$	$1/0.05^3$	$1.4 \quad 10^4$	14
4th	$485^4/1$	$1/0.05^4$	$3.5 \quad 10^5$	19

entries in F is used as a measure of demand on numerical precision. See Table 1. The ratio of the span of entries in F for the normalized vs. unnormalized cases is given as "Cost Ratio of Spans". This gives a measure for the reduction in precision that is achieved by normalizing. This reduction is then expressed as a number of bits.

The table shows that a significantly larger number of bits can be required to represent the ratios for the unnormalized entries. This unnecessary cost in precision is reflected in poorer conditions for F, reaching 6 orders of magnitude! (See Tables 2 and 3). PRIME has been calibrated to an accuracy of 1 part in 1500 (see Section 5.3). This necessitates ≈ 11 bits of precision for ranging, nominally. The burden of using non-normalized formulations is additive. Note that these costs in precision have been expressed as a number of bits in integer format. The actual numerical effects are more complicated, as these involve floating point operations.

A Singular Value Decomposition (SVD) [56] exposes the singular values of F, which can be used to compute the condition number [57]. This provides an excellent measure of the numerical sensitivity of the solution. The SVD approach decomposes F

$$F = UDV^T \tag{7}$$

where U and V are orthonormal and D is diagonal, containing the singular values of F. The

technique also permits "repairs" to be made to D, ensuring its invertibility [58]. This allows the calibration model

$$M = VD^{-1}U^{T}F \tag{8}$$

and the condition number

$$\kappa = d_{max}/d_{min} \tag{9}$$

to be easily computed. Here, d_{max} and d_{min} are the largest and smallest diagonal elements of D. F is considered to be ill-conditioned if $1/\kappa$ approaches the precision of floating point computations, for example, no less than 10^{-12} for double precision [58].

5.3 Evaluation of Calibration Model

Selecting an appropriate form for a calibration model effects important tradeoffs in acquisition speed and ranging accuracy. This selection process is closely related to that associated with the Two-Planes method of camera calibration [55]. These problems are quite similar because each involve a mapping from image coordinates to a plane of world coordinates.

Recommendations for appropriate forms of camera models [59] serve as a guide, as well as several metrics for model evaluation [29, 60] which have been incorporated and extended for PRIME. The metrics used for PRIME involve the computational burden during on-line evaluation and various measures of the quality of the model solution. These are:

1) Compute burden

2) Condition number

3) Span of residual errors

4) Autocorrelation of residual errors

5) Goodness-of-fit, based on chi-squared.

The compute burden is a direct result of the number of terms in the model. The condition number is found from Eq. 9. The residual error of the solution is computed by comparing the fitted height of the ith triplet, $z_i' = f(r,c)_i$ to the original height. Ideally $z_i - z_i' \rightarrow 0$, but in practice random measurement noise and higher order lens aberrations keep these from vanishing. Although low residual errors do not guarantee an accurate model, these values are still useful to examine, as a low magnitude is a necessary condition for accuracy. The span of the residual errors is described by computing the maximum, mean and standard deviation of the absolute residuals.

The purpose of calibration is to characterize the distortions in a mapping between two planes. Hence, a proper mapping and proper calibration procedure should result in a set of residual errors with a very low spatial dependence. That is, any uniform pattern or trend in the spatial arrangement of residuals should be very minimal. The residuals should appear as random entries. This random character can be described using the autocorrelation of the residuals.

As a final metric, a goodness-of-fit measure has been computed involving chi-squared and the degrees of freedom in the system of equations [58]. This involves the use of an estimate of the accuracy of z_i, which is noted in the tables. The fit quality, Q, should ideally be $Q \geq 0.1$ to consider a mapping as valid. In some cases values in the range $Q \geq 0.001$ are also deemed acceptable [58].

Tables 2 and 3 summarize the metrics found when calibrating PRIME. Table 4 gives the form of each model that was considered. When finding the Z model, 280 calibration triplets were used

Table 2: *Metrics for Z calibration model for PRIME. A total of 280 calibration triplets were used. A value of 0.005 inches was used for the accuracy of z_i.*

Type	Normalized	Mean Residual (in.)	Standard Deviation of Residual (in.)	Maximum Residual (in.)	Fit Quality	Autocorrelation (1 shift)	Condition Number
1	yes	0.071	0.038	0.137	0.00	0.9	10^1
2	yes	0.007	0.004	0.022	0.00	0.9	10^2
3	yes	0.003	0.003	0.015	0.99	0.7	10^3
4	yes	0.002	0.002	0.013	1.00	0.5	10^3
1	no	0.071	0.038	0.137	0.00	0.9	10^3
2	no	0.007	0.004	0.022	0.00	0.9	10^6
3	no	0.003	0.003	0.015	0.99	0.7	10^9
4	no	0.002	0.002	0.013	1.00	0.5	10^9

Table 3: *Metrics for X calibration model for PRIME. A total of 98 calibration triplets were used. A value of 0.010 inches was used for the accuracy of x_i.*

Type	Normalized	Mean Residual (in.)	Standard Deviation of Residual (in.)	Maximum Residual (in.)	Fit Quality	Autocorrelation (1 shift)	Condition Number
1	yes	0.093	0.059	0.240	0.00	0.8	10^1
2	yes	0.010	0.008	0.033	0.00	0.8	10^2
3	yes	0.009	0.007	0.031	0.03	0.8	10^3
4	yes	0.004	0.002	0.009	1.00	0.8	10^3
1	no	0.093	0.059	0.240	0.00	0.8	10^3
2	no	0.010	0.008	0.033	0.00	0.8	10^6
3	no	0.009	0.007	0.031	0.03	0.8	10^9
4	no	0.004	0.002	0.009	1.00	0.8	10^9

and the accuracy of the true heights, z_i, was estimated at 0.005 inches. For Y, 98 triplets and an accuracy of 0.010 inches for x_i were used.

In addition to the tabular summaries, images have also been generated that depict the residual errors. These provide a visualization of any spatial correlation in the errors and of the location of all the calibration triplets, in a single image. Figures 11 through 14 show images that depict

Table 4: *Form of calibration models.*

Type	Coefficients
1	$1\ r\ c$
2	$1\ r\ c\ r^2\ c^2\ rc$
3	$1\ r\ c\ r^2\ c^2\ rc\ r^3\ c^3$
4	$1\ r\ c\ r^2\ c^2\ rc\ r^2c\ rc^2\ r^3\ c^3$

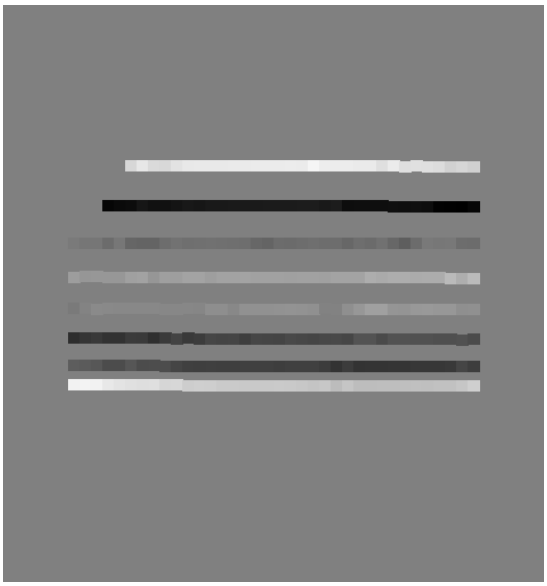

Figure 11: *Residual error of a 1st order (type 1) Z calibration model. Mean = 0.071 in. Gray levels of each block are scaled to span the extremes of the residual errors in this model (maximum = 0.137 in).*

residual errors for types 1 and 4 of the Z and X models, respectively. The images contain a series of small gray blocks, each of which coincides to the location of a calibration triplet and the corresponding residual error. The gray level of the block is scaled so that the data point possessing a minimum residual error is the darkest and the one possessing the largest error is the lightest. Note that a different gray scale mapping is used for each image.

As seen in Table 5, the mean acquisition time did not vary significantly between model types. This is primarily due to the compute burden associated with managing the datacube hardware and with finding the nominal position of the laser profile. Because of this, model selection was based primarily on accuracy and condition number. The normalized form of model type 4 was chosen for both X and Z.

Figure 12: *Residual error of a 3rd order (type 4) Z calibration model. Mean = 0.002 in. Gray levels of each block are scaled to span the extremes of the residual errors in this model (maximum = 0.013 in).*

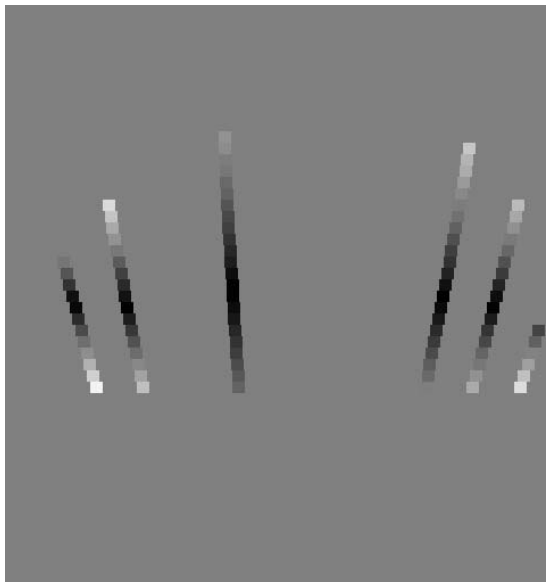

Figure 13: *Residual error of a 1st order (type 1) X calibration model. Mean = 0.093 in. Gray levels of each block are scaled to span the extremes of the residual errors in this model (maximum = 0.240 in).*

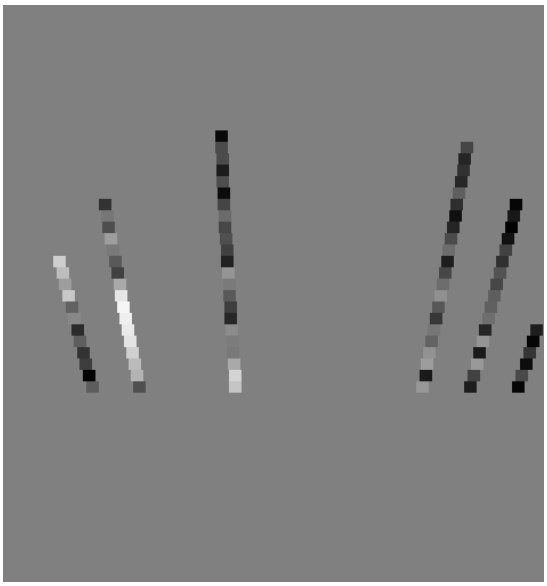

Figure 14: *Residual error of a 3rd order (type 4) Z calibration model. Mean = 0.004 in. Gray levels of each block are scaled to span the extremes of the residual errors in this model (maximum = 0.009 in).*

6 Sensitivity Analysis of Structured Light Sensors

The question of sensitivity is concerned with the relationship between the accuracy of range measurements to sources of error in the acquisition process. The geometry of the optical paths involved in ranging plays an important role in determining how errors during acquisition are amplified in range measurements.

A Monte Carlo analysis [61] can be used to perform a thorough study of measurement errors. This would involve simulations of the measurement process with representative levels of noise being introduced at each stage. This type of approach is particularly desirable when analyzing a SL sensor with dynamic geometry, as it can be used to study accuracies across the entire measurement space. For PRIME, a simpler approach has been taken.

Referring to Figure 15, variations in image coordinates, $d\phi$, produce a displacement of dw

at the standoff of the sensor. Because the laser profile is assumed to be roughly horizontal in each camera image and each column is analyzed individually during pixel-level processing, only the sensitivity of height variations to vertical image displacements is considered. (Actually in the center of the image $dy = dw$ for lateral image displacements.) The angle between the laser emission and the camera sighting determines the sensitivity of height errors. As seen in the figure, the amplifying factor is

$$\frac{dz}{dw} = \frac{1}{\sin\theta} \tag{10}$$

where θ is the angle between the laser emission and camera sighting. This is related to changes in height by $\sin\theta$.

In Fixed-Plane SL sensors this sensitivity factor varies in a continuous fashion across the laser plane because θ varies somewhat across the laser profile (with motion in and out of the paper in Figure 15). For PRIME, the sensitivity varies $1.5 <= \frac{dz}{dw} <= 1.6$. Being near 1, these values are relatively low. As sensitivity improves the degree of shadowing increases. The ranging geometry for PRIME was chosen to somewhat favor sensitivity versus shadowing. Freedom to increase shadowing was deemed acceptable because mostly convex objects with relatively slow rates of curvature were targeted for use with PRIME.

7 Performance Benchmarks and Ranging Experiments

Performance benchmarks are very important during sensor research. These give measures of speed and accuracy that are vital for both automated and manual interpretation of data. Example range data and scenes are presented in Figures 17 and 16, respectively. Images of the

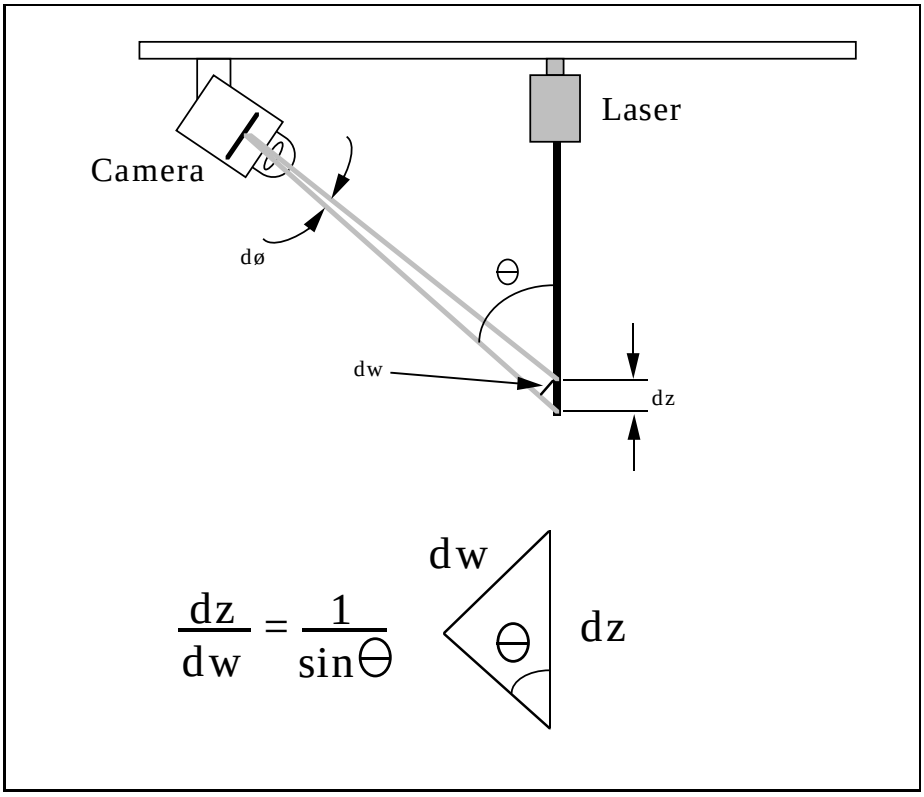

Figure 15: *The sensitivity dz/dw of a Fixed-Plane SL system is determined by the angle, θ, between the laser emission and the camera sighting. This describes the effect of the change in image coordinates with respect to changes in height.*

scenes were captured by the PRIME video camera with the IR bandpass filter removed. The laser appears as a bright line near the bottom of each image.

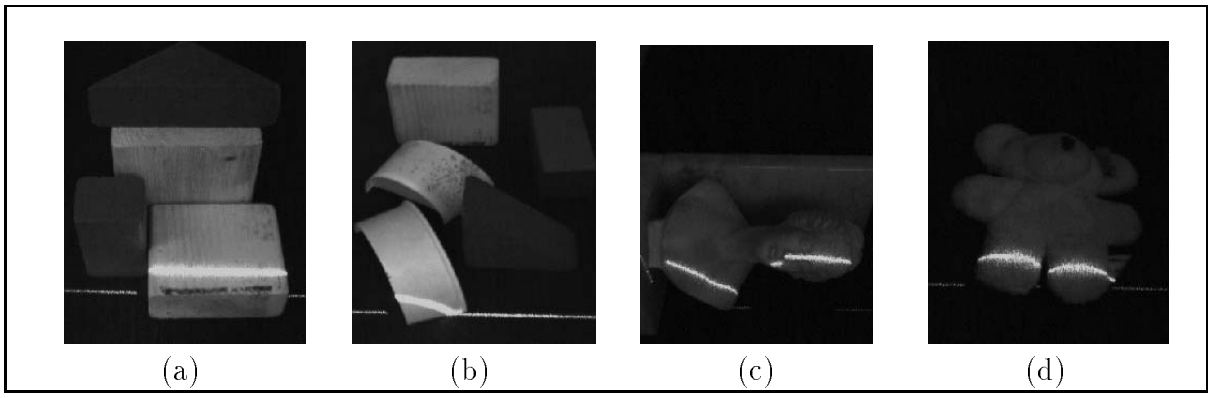

Figure 16: *Gray scale images captured from the video camera during the ranging process. The optical filter normally used during acquisition was removed for these images, to better illustrate the interaction between the laser plane and object surfaces. The laser appears as a bright line near the bottom of each image.*

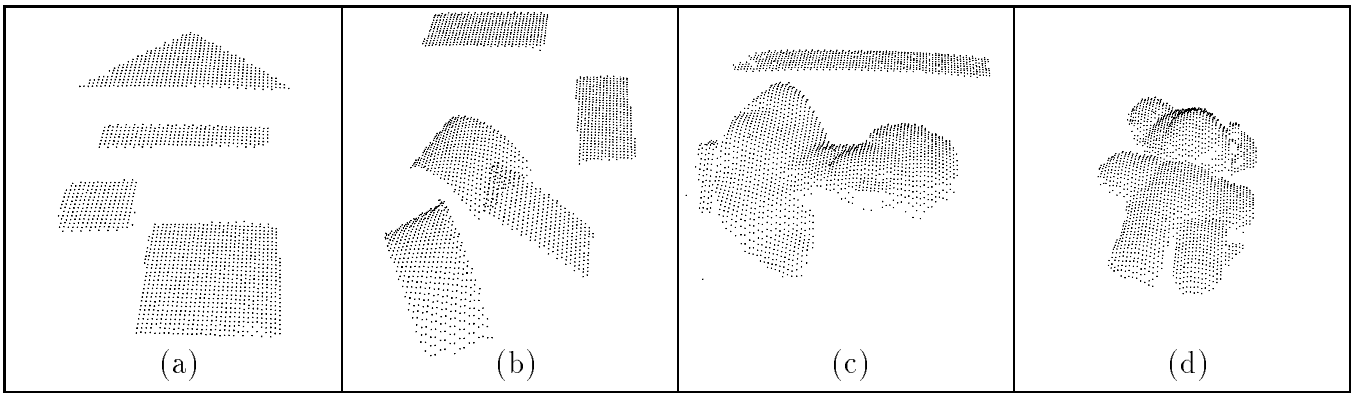

Figure 17: *Measurements captured by PRIME for objects with a variety of surface textures.*

A performance metric for active range sensors has been introduced by [24].

$$M = \frac{1}{T^{1/2}} \frac{(D_x D_y D_z)^{1/3}}{(\delta_x \delta_y \delta_z)^{1/3}} \tag{11}$$

where T is the point dwell time (sec/point), D_l is the depth of field in the l direction, and δ_l is the measurement uncertainty, also in the l direction. M is the rating. In [24] the dimensionality

of contributors, (x, y, z), to this expression is reduced for sensors that generate indexed arrays of range data. This is the case for PRIME in the X direction. Also with PRIME, the depth of field in the direction along the conveyer is somewhat arbitrary, as the conveyer is often run at varying speeds depending on application needs. As in [24], the rating has been reduced to

$$M = \frac{D_z}{\delta_z \sqrt{T}}. \tag{12}$$

To determine the accuracy of height measurements, a horizontal plate was scanned at a height other than those used during calibration. This plate and the ones used during calibration have a quoted flatness of 0.002 inch variation per foot. At most 6 inches of such a plate are used during calibration and testing. The mean absolute variation of height measurements taken from this type of plate is given by δ_z in Table 5.

To determine the point dwell time, a number of scans were made and the mean number of points per video frame was computed. Each type of calibration model was tested. The size of processed imagery was adjusted to maximize the amount of range data, while maintaining frame-rate throughput for each calibration model. The results of these performance benchmarks is given in Table 5. These benchmarks were also used in the model selection process. The T values in the table are somewhat conservative and are subject to several percent error, due to empirical nature with which they were determined.

These ratings compare well with those given in [24]. It should be noted that the speed benchmarks quoted here include the application of calibration models. Also, PRIME has been built from commercially available components. This makes for a system that is generally less expensive

Table 5: *PRIME performance measures as a function of the type of calibration model. Performance of the normalized vs. unnormalized forms was essentially identical, normalized models were used herein.*

Type	δ_z	T	M
1	0.005	0.24	51,500
2	0.004	0.25	70,200
3	0.004	0.25	70,200
4	0.003	0.25	83,900

and more easily maintained than a custom implementation.

8 Related Research and Concluding Remarks

Across the diverse spectrum of Machine Vision applications, the main objective is often the same: to extract useful information from image inputs. For tasks requiring 3-D information, Machine Vision techniques may be grouped into passive or active approaches. Active approaches, such as Structured Light, use specialized illumination sources to overcome the ambiguities associated with passive methods.

Despite the longevity of research in Structured Light sensing, a limited amount of published works focus on fundamental design and calibration issues. This chapter has included an introduction to the ranging process, discussions of design tradeoffs, calibration methods, and performance benchmarks. Structured Light ranging has some particularly interesting advantages that allow sensors to be customized for the specific requirements of an application. These sensors can be built "from the ground, up" to yield rugged and inexpensive ranging systems.

PRIME is a Structured Light sensor that has been designed to scan continuously moving objects. PRIME uses a plane of laser light that is mounted in a fixed geometry. The laser

illuminates scenes, allowing vertical profiles of range data to be acquired in real time.

Range sensors can be used to form the foundations of larger, end-to-end Machine Vision systems. This involves a number of processing steps. Figure 18 illustrates the components of a recognition system involving range data and graph-matching techniques. Many versions of similar diagrams have been reported [1, 2].

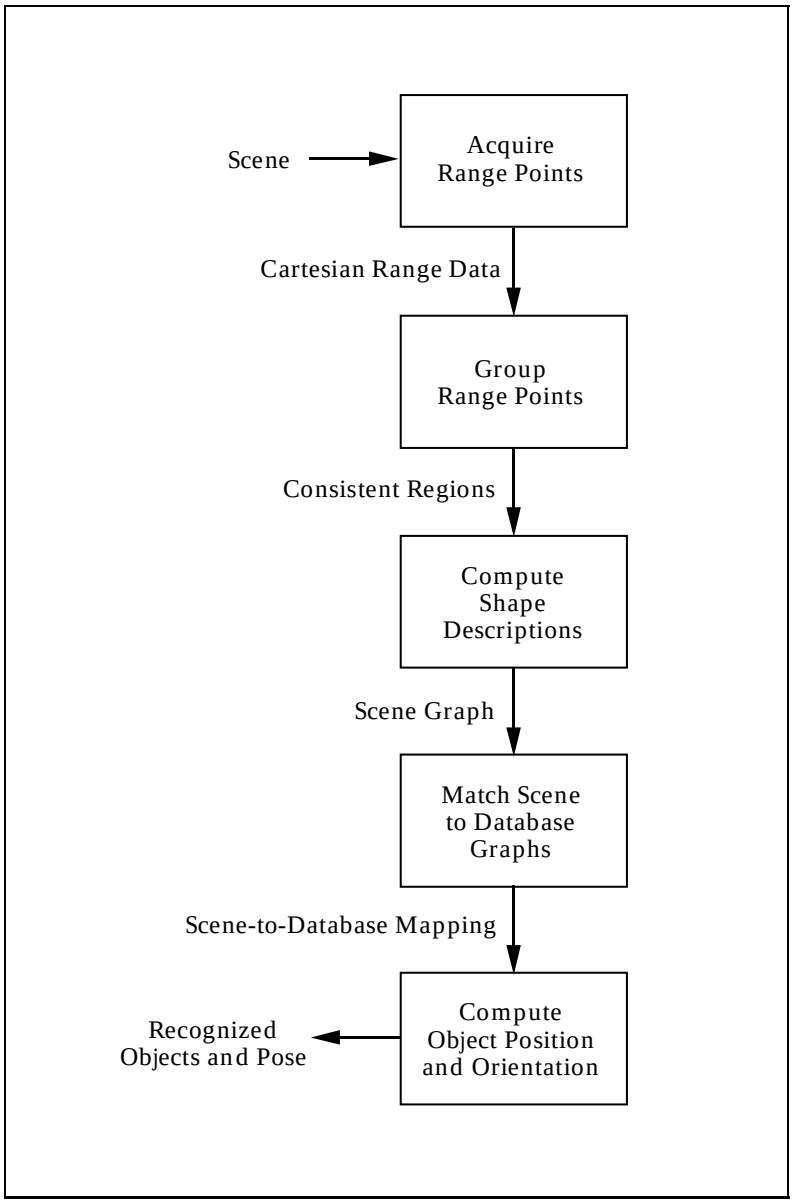

Figure 18: *Processing steps involved in an end-to-end object recognition system.*

"Acquire Range Points" generates Cartesian range points, as is the purpose of PRIME. Many range sensors are reviewed in [23, 24]. Range points have no inherent high-level meaning, they are simply individual, 3-D measurements in space. The first step in higher-level interpretation is to "Group Range Points" into consistent regions. This process is commonly referred to as segmentation and has been an active area of research for over 15 years [62, 63, 64]. Reviews are available in [65, 66]. A novel and real-time approach to range segmentation has been developed for use with PRIME [67, 68, 69]. Results from this process are illustrated in Figure 19.

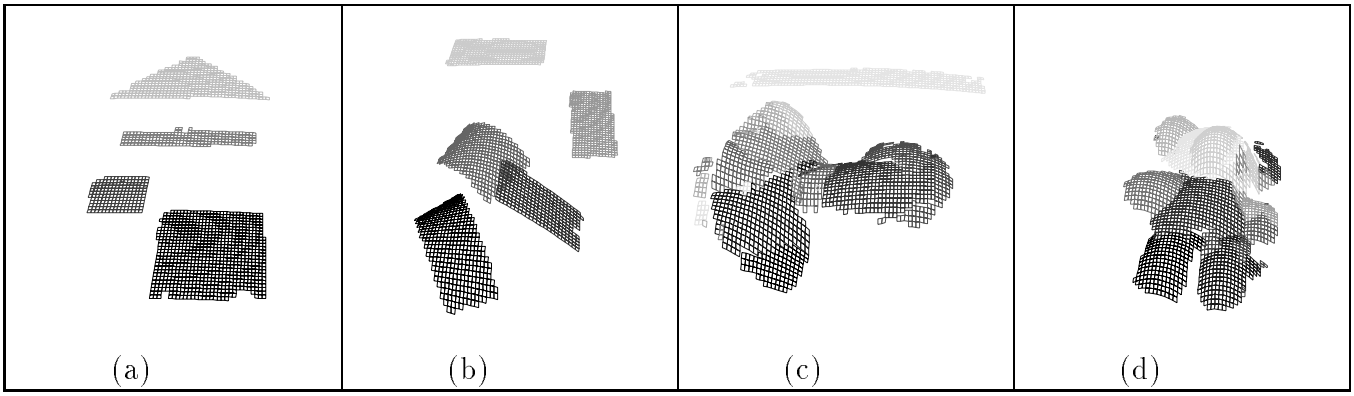

Figure 19: *Results of segmentation of range measurements take by PRIME. Each consistent region of range data is shown with a surface patch of a different gray tone. Both flat and curved patches are illustrated.*

After "Consistent Regions" have been formed, models are typically used to form a parameterized description of each region. Various types of generalized geometric models are common [70, 71, 72] as are approaches based on CAD models [73, 74]. Surveys may be found in [75, 76, 77]. This remains a very active area of research.

Graphs are typically used to describe the adjacency relationships of scene elements. Graph-based descriptions also provide a framework for matching schemes used to recognize objects [78, 79, 3]. An immense amount of research has also been pursued in the areas of recognition and object localization. In Figure 18 these are presented as two steps. Some implementations keep these

operations distinct [80, 81, 82, 83], in many other approaches recognition and pose calculations are accomplished in a coupled process [84, 85, 86, 48, 38].

A relatively low number of end-to-end recognition systems have been reported, compared to the very large amount of effort on system components. Complete systems can be found in [71, 86, 48, 38, 73, 77]. PRIME is being integrated into an end-to-end recognition system. This includes the novel range segmentation strategy in [67, 69] and a recognition strategy based on graph-matching [87]. Applications in the Active Vision discipline [17, 88, 20, 3] and in geographically-distributed real-time manufacturing are being targetted with this system.

References

[1] M. M. Trivedi and A. Rosenfeld, "On making computers see," *IEEE Trans. on System, Man, and Cybernetics*, vol. 19, no. 6, pp. 1333–1335, 1989.

[2] A. Rosenfeld, *Advances in Computers*, vol. 27, ch. Computer Vision. Boston: Academic Press, 1988.

[3] M. M. Trivedi and C. Chen, *Advances in Computers*, vol. 32, ch. Sensor-Driven Intelligent Robotics, pp. 105–148. Boston: Academic Press, 1988.

[4] M. M. Trivedi, *Encyclopedia of Science and Technology*, ch. Intelligent Robotic Systems, pp. 226–229. New York: McGraw-Hill, 1994.

[5] C. Thorpe, M. Hebert, T. Kanade, and S. A. Shafer, "Visiosn and navigation for the carnegie-mellon NAVLAB," *IEEE Trans. on Pattern Anal. Machine Intell.*, pp. 362–373, May 1988.

[6] C. A. Harlow, M. M. Trivedi, R. W. Conners, and D. Phillips, "Scene analysis of high resolution aerial scenes," *Optical Engineering*, pp. 347–355, March 1986.

[7] P. H. Eichel, E. J. Delp, K. Koral, and A. J. Buda, "A method for fully automatic definition of coronary arterial edges from cineangiograms," *IEEE Trans. Medical Imaging*, pp. 313–320, December 1988.

[8] L. S. Hibbard, J. S. McGlone, D. W. Davis, and R. A. Hawkins, "Three dimensional representation and analysis of brain energy metabolism," *Science*, pp. 1641–1646, 1987.

[9] A. M. Darwish and A. K. Jain, "A rule-based approach for visual pattern inspection," *IEEE Trans. on Pattern Anal. Machine Intell.*, pp. 56–68, January 1988.

[10] R. Kasturi, S. T. Bow, W. El-Masri, J. Shah, J. R. Gattiker, and U. B. Mokate, "A system for interpretation of line drawings," *IEEE Trans. on Pattern Anal. Machine Intell.*, pp. 978–992, October 1990.

[11] A. Okakzaki, T. Kondo, K. Mori, S. Tsunekawa, and E. Kawamoto, "An automatic circuit diagram reader with loop-structure-based symbol recognition," *IEEE Trans. on Pattern Anal. Machine Intell.*, pp. 331–341, May 1988.

[12] D. Marr, *Vision*. San Fransico, CA: Freeman, 1982.

[13] M. Okutomi, G. Medioni, and T. Kanade, "A multiple baseline stereo," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 4, pp. 353–363, 1993.

[14] S. B. Marapane and M. M. Trivedi, "Region-based stereo analysis for robotic applications," *IEEE Trans. on System, Man, and Cybernetics*, vol. SMC-19, pp. 1447–1464, November 1989.

[15] S. B. Marapane and M. M. Trivedi, "An active vision system for multi-primitive hierarchical stereo analysis and multi-cue depth extraction," in *Proceedings of the Sensor Fusion for Aerospace Systems Conference.*, (Orlando), SPIE, April 1993.

[16] S. B. Marapane and M. M. Trivedi, "Multi-primitive hierarchical (MPH) stereo analysis," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 16, pp. 227–240, March 1995.

[17] S. B. Marapane and M. M. Trivedi, "Experiments in active vision with real and virtual robot heads," *International Journal of Applied Intelligence, Special issue on Sensor Fusion*, vol. 5, pp. 237–250, July 1995.

[18] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical Engineering*, pp. 139–144, January/February 1980.

[19] D. Blostein and N. Ahuja, "Shape from texture: Integrating texture element extraction and surface estimation," *IEEE Trans. on Pattern Anal. Machine Intell.*, pp. 1233–1251, December 1989.

[20] A. K. Dalmia and M. M. Trivedi, "Real-time depth extraction using image streams acquired by a single camera," *Computer Vision and Image Understanding*, To appear 1995.

[21] A. K. Dalmia and M. M. Trivedi, "3-d structure extraction from image streams: A review," *IEEE Trans. on Pattern Anal. Machine Intell.*, To appear 1996.

[22] A. K. Dalmia and M. M. Trivedi, "Integrating depth from motion and binocular stereo approaches," *Machine Vision and Applications*, To appear 1996.

[23] R. A. Jarvis, "A perspective on range finding techniques for computer vision," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 5, no. 2, pp. 122–139, 1983.

[24] P. J. Besl, "Active, optical range imaging sensors," *Machine Vision and Applications*, vol. 1, pp. 127–152, 1988.

[25] G. Stockman and G. Hu, "Sensing 3-d surface patches using a projected grid," in *Proceedings Computer Vision and Pattern Recognition Conf.*, pp. 602–607, 1986.

[26] M. R. Ward, D. P. Rheaume, S. W. Holland, and J. H. Dunseth, "Production plant consight installations," Tech. Rep. GMR-4156, General Motors Research Laboratories, Warren, Michigan, August, 1982.

[27] H. P. Gadagkar, M. M. Trivedi, and T. N. Lassiter, "Versatile multi-modal system for surface profile measurements using a wrist-mounted laser device," in *Proceedings of the Sensor Fusion V Conference*, (Boston), pp. 466–474, SPIE, Nov. 1992.

[28] A. M. McIvor and R. J. Valkenburg, "Calibrating a structured light system," Tech. Rep. 362, Industrial Research Limited, February 1995.

[29] F. W. DePiero and R. L. Kress, "Design and in situ calibration of a structured light sensor," in *Proceedings of Intl. Conf. on Intelligent Robotics and Systems*, (Pittsburgh, PA), pp. 513–518, IEEE/RSJ, August 5-9 1995.

[30] B. L. Burks, F. W. DePiero, J. C. Rowe, C. B. Selleck, and D. L. Jacoboski, "Final results of the application of a structured light source for surface mapping of the fernald k-65 silos," in *Proceedings of The American Nuclear Society*, (Boston), ANS, June 1992.

[31] S. Kweon, R. Hoffman, and E. Krotkov, "Experimental characterization of the perceptron laser rangefinder," Tech. Rep. CMU-RI-TR-91-1, The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, 1991.

[32] S. W. Holland, L. Rossol, and M. R. Ward, *Computer Vision and Sensor-Based Robotics*. New York: Plenum Press, 1978.

[33] T. Kanade, A. Gruss, and L. R. Carley, "A very fast VLSI rangefinder," in *Proceedings of the 1991 IEEE Intl. Conf. on Robotics and Automation*, (Sacramento, CA), pp. 1322–1329, IEEE, April 1991.

[34] F. W. DePiero and R. L. Kress, "Camera calibration in a hazardous environment performed in situ with automated analysis and verification," in *Proceedings of ANS Fifth Topical Meeting on Robotics and Remote Systems*, (Knoxville, TN), ANS, April 1993.

[35] K. L. Boyer and A. C. Kak, "Color encoded structured light for rapid active ranging," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 9, no. 1, pp. 14–28, 1987.

[36] Y. Shirai and M. Suwa, "Recognition of polyhedrals with a range finder," in *Proceedings of the 2nd Int. Joint Conf. on Artificial Intell.*, (London), pp. 80–87, Sept. 1971.

[37] G. Agin and T. Binford, "Computer description of curved objects," in *Proceedings of the 3rd Int. Joint Conf. on Artificial Intell.*, pp. 629–640, 1973.

[38] M. Oshima and Y. Shirai, "Object recognition using three dimensional information," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 5, no. 4, pp. 353–361, 1983.

[39] M. Oshima and Y. Shirai, "A scene description method using three dimensional information," *Pattern Recognition*, vol. 11, pp. 9–17, 1984.

[40] S. Tada, A. Gruss, and T. Kanade, "Cmu very fast range-imaging system," Tech. Rep. CMU-CS-93-179, School of Computer Science, 1993.

[41] B. L. Burks, F. W. DePiero, J. C. Rowe, C. B. Selleck, D. L. Jacoboski, and R. Markus, "Generation of 3 surface maps in waste storage silos using a structured light source," in *Proceedings of The Space Operations, Applications, and Research Symposium*, (Houston), NASA, 1991.

[42] M. Minou, *Theoretical and experimental studies on basic relations between real world and pictoral patterns and their generating constraints.* PhD dissertation, Kyoto University, Department of Information Science, Nov. 1982.

[43] M. D. Altscheuler, B. R. Altscheuler, and J. Toboada, "Laser electro-optic system for rapid 3-d topographic mapping of surfaces," *Optical Engineering*, vol. 20, no. 6, pp. 953–961, 1981.

[44] S. Inokuchi, K. Sato, and F. Matsuda, "Range imaging system for 3-d object recognition," in *Proceedings of the 7th International Conf. on Pattern Recog.*, pp. 806–808, 1984.

[45] K. K. Yeung and P. D. Lawrence, "A low-cost 3d vision system using space-encoded spot projections," in *Proceedings of The SPIE Conf. on Optics, Illumination and Image Sensing for Machine Vision*, pp. 160–172, SPIE, 1986.

[46] R. J. Popplestone, C. M. Brown, A. P. Ambler, and G. F. Crawford, "Forming models of plane-and-cylinder faceted bodies from light stripes," in *Proc. 4th Int. Joint Conf. on Artificial Intell.*, pp. 664–668, 1975.

[47] G. Bickel, G. Hausler, and M. Maul, "Triangulation with expanded range of depth," *Optical Engineering*, vol. 24, no. 6, pp. 975–979, 1985.

[48] O. D. Faugeras and M. Hebert, "The representation, recognition and locating of 3d shapes from range data," *International Journal of Robotics Research*, vol. 5, no. 3, pp. 27–52, 1986.

[49] M. Rioux, "Laser range finder based on synchronized scanners," *Applied Optics*, vol. 23, no. 21, pp. 3837–3855, 1984.

[50] B. K. P. Horn, *Robot Vision.* Cambridge, MA: McGraw-Hill, 1984.

[51] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans ASME J. Basic Eng.*, vol. 83, pp. 35–45, 1960.

[52] J. A. Borrie, *Stochastic Systems for Engineers: Modeling, Estimation and Control.* New York: Prentice Hall, 1992.

[53] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association.* New York: Academic Press, 1989.

[54] G. R. Cooper and C. D. McGillem, *Probabalistic Methods of Signal and System Analysis.* Orlando, FL: Harcourt Brace Jovanovich, 1986.

[55] A. Isaguirre, P. Pu, and J. Summers, "A new development in camera calibration: Calibrating a pair of mobile cameras," tech. rep., Department of Computer and Information Science, GRASP Laboratory, University of Pennsylvania, Philadelphia, 1985.

[56] G. H. Golub and C. F. V. Loan, *Matrix Computations.* Baltimore: John Hopkins University Press, 2nd ed., 1989.

[57] G. Strang, *Linear Algebra and its Applications.* New York: Academic Press, 2nd ed., 1980.

[58] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C.* Cambridge, New York: Cambridge University Press, 1988.

[59] G.-Q. Wei and S. D. Ma, "Two plane camera calibration: A unified model," tech. rep., National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Bejing, 1991.

[60] S. M. Thayer and M. M. Trivedi, "Residual uncertainty in three-dimensional reconstruction using two-planes calibration and stereo methods," *Pattern Recognition*, vol. 28, no. 7, pp. 1073–1082, 1995.

[61] I. M. Sobol, *A Primer for the Monte Carlo Method.* Boca Raton: CRC Press, 1994.

[62] R. Duda and P. Hart, *Pattern Classification and Scene Analysis.* New York: John Wiley and Sons, 1973.

[63] R. C. Bolles and M. A. Fischler, "A ransac-based approach to model fitting and its application to finding cylinders in range data," in *Proceedings of the 7th International Joint Conf. on Artificial Intelligence*, (Vancouver, B.C., Canada), pp. 637–643, IJCAI, Aug 24-28 1981.

[64] K. L. Boyer, M. J. Mirza, and G. Ganguly, "The robust sequential estimator: A general approach and its application to surface organization in range data," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 16, no. 10, pp. 987–1001, 1994.

[65] P. J. Besl and R. C. Jain, "Three-dimensional object recognition," *ACM Computing Surveys*, vol. 17, no. 1, 1985.

[66] F. W. DePiero and M. M. Trivedi, "The evolution of adaptive techniques for range image segmentation," *Submitted to IEEE Trans. Pattern Analysis and Machine Intelligence*, 1996.

[67] F. W. DePiero and M. M. Trivedi, "Real-time range image acquisition and segmentation using adaptive kernels and kalman filtering," in *Submitted to Intl. Conf. on Robotics and Automation*, (Minneapolis, MN), IEEE, April 22-28 1996.

[68] F. W. DePiero and M. M. Trivedi, "A synergistic and real-time approach to range image acquisition and segmentation using a profile-based technique," in *Submitted to Computer Vision and Pattern Recognition*, (San Francisco, CA), IEEE Computer Society, June 18-20 1996.

[69] F. W. DePiero and M. M. Trivedi, "Range profile tracking (RPT): A profile-based approach for robust and real-time range image segmentation," *Submitted to IEEE Pattern Analysis and Machine Intelligence.*

[70] B. K. P. Horn, "Extended gaussian image," *Proceedings of IEEE*, vol. 61, no. 3, pp. 1671–1686, 1984.

[71] R. A. Brooks, "Model-based three-dimensional interpretations of two-dimensional images," *IEEE Trans. on Pattern Anal. Machine Intell.*, pp. 140–150, March 1983.

[72] W. E. L. Grimson and T. Lozano-Perez, "Model-based recognition and localization from sparce range or tactile data," *International Journal of Robotics Research*, vol. 3, pp. 3–35, Fall 1983.

[73] R. Bolles and P. Horaud, "3DPO: A three-dimensional part orientation system," *International Journal of Robotics Research*, vol. 5, no. 3, pp. 3–26, 1986.

[74] P. J. Flynn and A. K. Jain, "Cad-based computer vision: From cad models to relational graphs," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 13, no. 2, pp. 114–132, 1991.

[75] R. T. Chin and C. R. Dyer, "Model-based recognition in robot vision," *ACM Computing Surveys*, vol. 18, no. 1, pp. 67–108, 1986.

[76] T. O. Binford, "Survey of model-based image analysis systems," *International Journal of Robotics Research*, pp. 18–64, Spring 1982.

[77] T.-J. Fan, G. Medioni, and R. Nevatia, "Recognizing 3-d objects using surface descriptions," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 11, no. 11, pp. 1140–1157, 1989.

[78] R. Nevatia, *Machine Perception*. Englewood Cliffs, NJ: Prentice-Hall, 1982.

[79] D. H. Ballard and C. M. Brown, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 1982.

[80] A. P. Ambler, H. G. Barrow, C. M. Brown, R. M. Burstall, and R. J. Popplestone, "A versatile computer-controlled assembly system," *IJCAI*, pp. 298–307, 1973.

[81] W.-Y. Kim and A. C. Kak, "3-d object recognition using bipartite matching embedded in discrete relaxation," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 13, no. 3, pp. 224–251, 1991.

[82] M. D. Wheeler and K. Ikeuchi, "Sensor modeling, markov random fields and robust localization for recognizing partially occluded objects," in *Proceedings of The Image Understanding Workshop*, (Wash. D. C.,), DARPA, April 1993.

[83] O. D. Faugeras, *Three-Dimensional Computer Vision, A Geometric Viewpoint*. London: MIT Press, 1993.

[84] G. Stockman, "Object recognition and localization via pose clustering," *Computer Vision, Graphics and Image Processing*, pp. 361–387, June 1987.

[85] C. R. Bidlack and M. M. Trivedi, "Geometric model based object recognition and localization robotic manipulation tasks," in *Applications of Artificial Intelligence IX Conf.*, (Orlando), pp. 270–280, SPIE, April 1991.

[86] C. H. Chen and A. C. Kak, "A robot vision system for recognizing 3-d objects in low-order polynomial time," *IEEE Trans. on System, Man, and Cybernetics*, vol. 19, no. 6, pp. 1535–1563, 1989.

[87] F. W. DePiero, M. M. Trivedi, and S. Serbin, "Graph matching using a direct classification of node attendance," *Accepted by Pattern Recognition Journal*, 1996.

[88] H. P. Gadagkar and M. M. Trivedi, "An integrated system for active exploration using contact and non-contact sensors," in *Proceedings of The IEEE IROS '92 Conference*, (Raleigh, NC), IEEE, July 1992.