Passive range sensors rely only on intensity images to reconstruct depth (e.g., stereopsis, discussed in Chapter 7).

Passive range sensors are the subject of Chapters 7, 8, and 9, and are not discussed further here. Active range sensors exploit a variety of physical principles; examples are radars and sonars, Moiré interferometry, focusing, and triangulation. Here, we sketch the first three, and concentrate on the latter in greater detail.

Radars and Sonars. The basic principle of these sensors is to emit a short electromagnetic or acoustic wave, or *pulse*, and detect the return (echo) reflected from surrounding surfaces. Distance is obtained as a function of the time taken by the wave to hit a surface and come back, called *time of flight*, which is measured directly. By sweeping such a sensor across the target scene, a full range image can be acquired. Different principles are used in imaging laser radars; for instance, such sensors can emit an amplitude-modulated laser beam and measure the phase difference between the transmitted and received signals.

Moiré Interferometry. A Moiré interference pattern is created when two gratings with regularly spaced patterns (e.g., lines) are superimposed on each other. Moiré sensors project such gratings onto surfaces, and measure the phase differences of the observed interference pattern. Distance is a function of such phase difference. Notice that such sensors can recover absolute distance only if the distance of one reference point is known; otherwise, only relative distances between scene points are obtained (which is desirable for inspection).

Active Focusing/Defocusing. These methods infer range from two or more images of the same scene, acquired under varying focus settings. For instance, *shape-from-focus* sensors vary the focus of a motorized lens continuously, and measure the amount of blur for each focus value. Once determined the best focused image, a model linking focus values and distance yields the distance. In *shape-from-defocus*, the blur-focus model is fitted to two images only to estimate distance.

In the following section, we concentrate on *triangulation-based* range sensors. The main reason for this choice is that they are based on intensity cameras, so we can exploit everything we know on intensity imaging. Moreover, such sensors can give accurate and dense 3-D coordinate maps, are easy to understand and build (as long as limited accuracy is acceptable), and are commonly found in applications.

2.5.3 Active Triangulation

We start by discussing the basic principle of active triangulation. Then, we discuss a simple sensor, and how to evaluate its performance. As we do not know yet how to calibrate intensity cameras, nor how to detect image features, you will be able to implement the algorithms in this section only after reading Chapters 4 and 5.

The basic geometry for an active triangulation system is shown in Figure 2.15. A light projector is placed at a distance b (called *baseline*) from the center of projection of

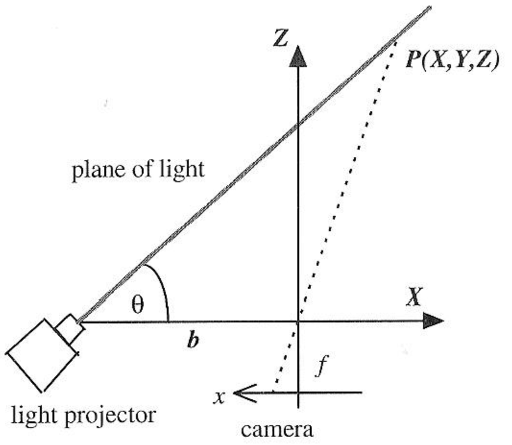

Figure 2.15 The basic geometry of active, optical triangulation (planar XZ view). The Y and y axes are perpendicular to the plane of the figure.

a pinhole camera.[9] The center of projection is the origin of the reference frame XYZ, in which all the sensor's measurements are expressed. The Z axis and the camera's optical axis coincide. The y and Y, and x and X axes are respectively parallel but point in opposite directions. Let f be the focal length. The projector emits a plane of light perpendicular to the plane XZ and forming a controlled angle, θ, with the XY plane. The Y axis is parallel to the plane of light and perpendicular to the page, so that only the profile of the plane of light is shown. The intersection of the plane of light with the scene surfaces is a planar curve called the *stripe*, which is observed by the camera. In this setup, the coordinates of a stripe point $\mathbf{P} = [X, Y, Z]^\top$ are given by

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{b}{f \cot\theta - x} \begin{pmatrix} x \\ y \\ f \end{pmatrix} \tag{2.24}$$

☞ The focal length and the other intrinsic parameters of the intensity camera can be calibrated with the same procedures to be used for intensity cameras (Chapter 6).

Applying this equation to all the visible stripe points, we obtain the 3-D profile of the surface points under the stripe (a cross-section of the surface). We can acquire multiple, adjacent profiles by advancing the object under the stripe, or sweeping the stripe across the object, and repeat the computation for each relative position of stripe and object. The sequence of all profiles is a full range image of the scene.

---

[9] Notice that, in Figure 2.15, the center of projection is *in front* of the screen, not behind, and is not the origin of the camera frame. This does not alter the geometry of image formation. Why?.

In order to measure (x, y), we must identify the stripe points in the image. To facilitate this task, we try to make the stripe stand out in the image. We can do this by projecting laser light, which makes the stripe brighter than the rest of the image; or we can project a black line onto a matte white or light grey object, so that the only really dark image points are the stripe's. Both solutions are popular but have drawbacks. In the former case, concavities on shiny surfaces may create reflections that confuse the stripe detection; in the latter, stripe location may be confused by shadows, marks and dark patches. In both cases, no range data can be obtained where the stripe is invisible to the camera because of occlusions. Sensors based on laser light are called *3-D laser scanners*, and are found very frequently in applications. A real sensor, modelled closely after the basic geometry in Figure 2.15, is shown in Figure 2.16.

☞ To limit occlusions one often uses two or more cameras, so that the stripe is nearly always visible from at least one camera.

2.5.4 A Simple Sensor

In order to use (2.24) we must calibrate f, b and θ. Although it is not difficult to devise a complete calibration procedure based on the projection equations and the geometry of Figure 2.17, we present here a simple and efficient method, called *direct calibration*, which does not require any equations at all. Altogether we shall describe a small but complete range sensor, how to calibrate it, and how to use it for measuring range profiles of 3-D objects. The algorithms require knowledge of some simple image processing operations, that you will be able to implement after going through the next three chapters.

The direct calibration procedure builds a lookup table (LUT) linking image and 3-D coordinates. Notice that this is possible because a one-to-one correspondence exists between image and 3-D coordinates, thanks to the fact that the stripe points are constrained to lie in the plane of light. The LUT is built by measuring the image coordinates of a grid of known 3-D points, and recording both image and world coordinates for each point; the depth values of all other visible points are obtained by interpolation.

The procedure uses a few rectangular blocks of known heights δ (Figure 2.17). One block (call it G) must have a number (say n) of parallel, rectangular grooves. We assume the image size (in pixels) is $x_{max} \times y_{max}$.

Algorithm RANGE_CAL

Set up the system and reference frame as in Figure 2.17. With no object in the scene, the vertical stripe falls on $Z = 0$ (background plane) and should be imaged near $y = y_{max} - 1$.

1. Place block G under the stripe, with grooves perpendicular to the stripe plane. Ensure the stripe appears parallel to x (constant y).
2. Acquire an image of the stripe falling on G. Find the y coordinates of the stripe points falling on G's higher surface (i.e., not in the groove) by scanning the image columns.
3. Compute the coordinates $[x_i, y_Z]^\top$, $i = 1, \ldots, n$, of the centers of the stripe segments on G's top surface, by taking the centers of the segments in the scanline $y = y_Z$. Enter each image point $[x_i, y_Z]^\top$ and its corresponding 3-D points $[X, Z]^\top$ (known) into a table T.

4. Put another block under G, raising G's top surface by δ. Ensure that the conditions of step 1 still apply. Be careful not to move the XYZ reference frame.
5. Repeat steps 2, 3, 4 until G's top surface is imaged near $y = 0$.
6. Convert T into a 2-D lookup table L, indexed by image coordinates $[x, y]^\top$, with x between 0 and $x_{max} - 1$, and y between 0 and $y_{max} - 1$, and returning $[X, Z]^\top$. To associate values to the pixels not measured directly, interpolate linearly using the four nearest neighbors.

The output is a LUT linking coordinates of image points and coordinates of scene points.

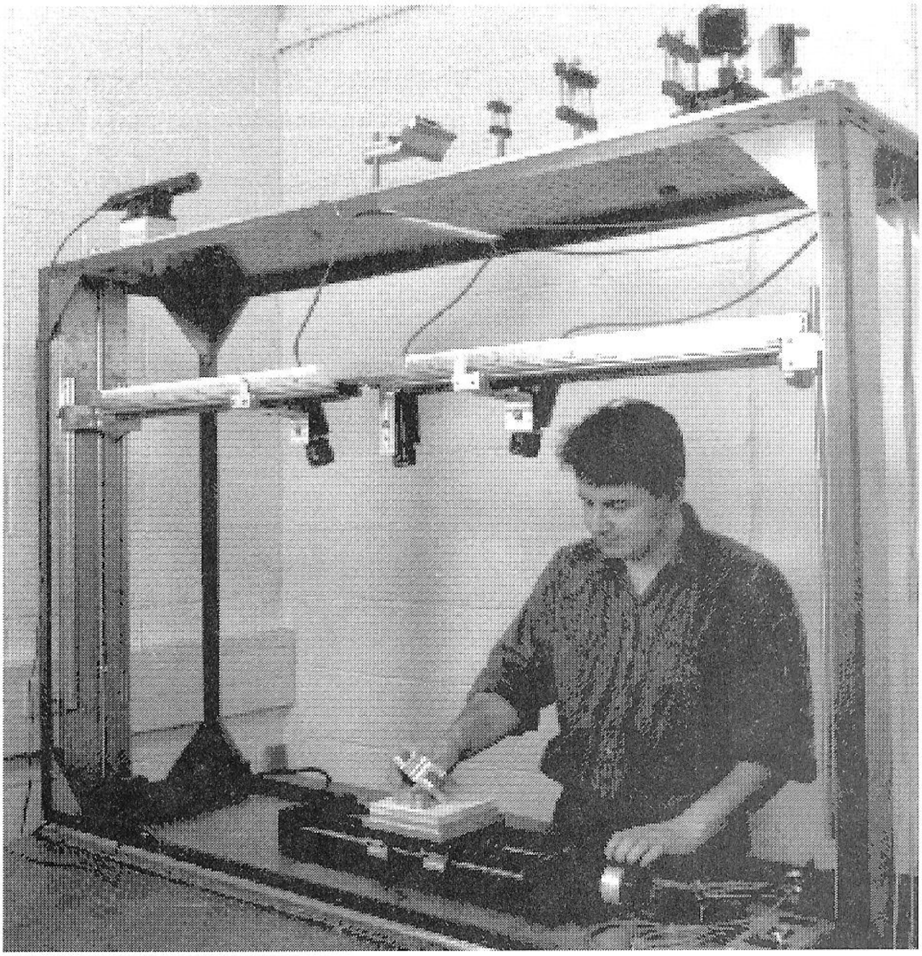

Figure 2.16 A real 3-D triangulation system, developed at Heriot-Watt University by A. M. Wallace and coworkers. Notice the laser source (top left), which generates a laser beam; the optical components forming the plane of laser light (top middle and left); the cameras; and the motorized platform (bottom middle) supporting the object and sweeping it through the stationary plane of light.

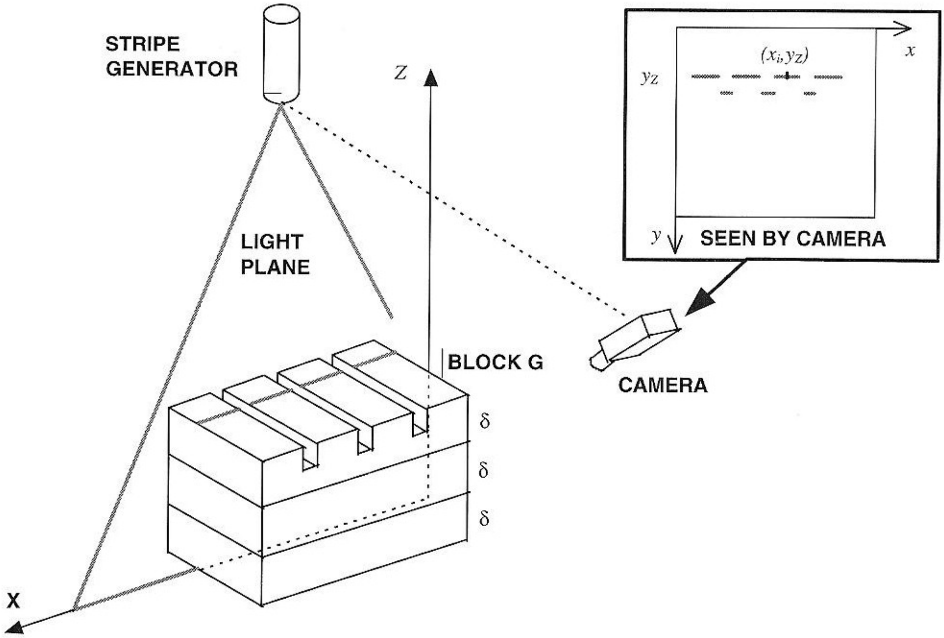

Figure 2.17 Setup for direct calibration of a simple profile range sensor.

And here is how to use L to acquire a range profile.

Algorithm RANGE_ACQ

The input is the LUT, L, built by RANGE_CAL.

1. Put an object under the stripe and acquire an image of the stripe falling on G.
2. Compute the image coordinates $[x, y]^\top$ of the stripe points by scanning each image column.
3. Index L using the image coordinates (x, y) of the stripe point, to obtain range points $[X, Z]^\top$.

The output is the set of 3-D coordinates corresponding to the stripe points imaged.

Notice that the numbers computed by such a sensor grow from the background plane ($Z = 0$), not from the camera.

☞ When a new block is added to the calibration scene, the stripe should move up by at least one or two pixels; if not, the calibration will not discriminate between Z levels. Be sure to use the *same* code for peak location in RANGE_ACQ *and* RANGE_CAL! The more sparse the calibration grid, the less accurate the range values obtained by interpolation in L.

When is a range sensor better than another for a given application? The following list of parameters is a basis for characterizing and comparing range sensors. Most parameters apply for non-triangulation sensors too.

Basic Parameters of Range Sensors

Workspace: the volume of space in which range data can be collected.

Stand-off distance: the approximate distance between the sensor and the workspace.

Depth of field: the depth of the workspace (along Z).

Accuracy: statistical variations of repeated measurements of a known true value (ground truth). Accuracy specifications should include at least the mean absolute error, the RMS error, and the maximum absolute error over N measures of a same object, with $N \gg 1$.

Resolution or precision: the smallest change in range that the sensor can measure or represent.

Speed: the number of range points measured per second.

Size and weight: important in some applications (e.g., only small sensors can be fitted on a robot arm).

☞ It is often difficult to know the *actual* accuracy of a sensor without carrying out your own measurements. Accuracy figures are sometimes reported without specifying to which error they refer to (e.g., RMS, absolute mean, maximum), and often omitting the experimental conditions and the optical properties of the surfaces used.

2.6 Summary

After working through this chapter you should be able to:

- ❑ explain how digital images are formed, represented and acquired
- ❑ estimate experimentally the noise introduced in an image by an acquisition system
- ❑ explain the concept of intrinsic and extrinsic parameters, the most common models of intensity cameras, and their applicability
- ❑ design (but not yet implement) an algorithm for calibrating and using a complete range sensor based on direct calibration

2.7 Further Readings

It is hard to find more on the content of this chapter on just one book. As a result if you want to know more you must be willing to do some bibliographic search. A readable account of basic optics can be found in the Feynman's Lecture on Physics [4]. A classic on the subject and beyond is the Born and Wolf [3] . The Born and Wolf also covers topics like image formation and spatial frequency filtering (though it is not always simple to go through). Our derivation of (2.13) is based on Horn and Sjoberg [6]. Horn [5] gives an extensive treatment of surface reflectance models. Of the many, very good textbooks on signal theory, our favorite is the Oppenheim, Willsky, and Young [11]. The discussion on camera models via the projection matrix is based on the appendix of Mundy and Zisserman's book *Geometric Invariants in Computer Vision* [9].

Our discussion of range sensors is largely based on Besl [1], which is a very good introduction to the principles, types and evaluation of range sensors. A recent,