CHAPTER

25 Image Acquisition

In vision, everything depends on image acquisition, and in image acquisition, everything depends on illumination. Naturally, robust algorithms can be designed to largely overcome any problems of inadequacy on these fronts. On the other hand, care with acquisition often means that simpler, more reliable algorithms can be produced. This chapter considers these important aspects of vision system design.

Look out for:

- lighting effects, reflectance, and the appearance of highlights and shadows.
- the value of soft or diffuse lighting.
- how lighting can systematically be made uniform by use of several point or line sources.
- the types of camera and sensor that are commonly available.
- the sampling theorem and its implications.

The advent of solid-state cameras and widely available frame-grabbing devices has made one part of image acquisition straightforward: yet the other aspect—that of providing suitable illumination—is still rather a black art. However, the methods described here demonstrate that uniform illumination is subject to design rather than *ad hoc* experimentation.

This work described in this chapter necessarily provides underpinning for all practical vision systems, except perhaps those involving X-rays or other modalities such as ultrasonic imaging.

25.1 INTRODUCTION

When implementing a vision system, nothing is more important than image acquisition. Any deficiencies of the original images can cause great problems

Computer and Machine Vision.

with image analysis and interpretation. An obvious example is that of lack of detail due to insufficient contrast or poor focusing of the camera: this can have the effect at best that the dimensions of objects will not be accurately measurable from the images, and at worst that the objects will not even be recognizable, so the purpose of vision cannot be fulfilled. This chapter examines the problems of image acquisition.

Before proceeding, it is as well to note that vision algorithms are of use in a variety of areas where visual pictures are not directly input. For example, vision techniques (image processing, image analysis, recognition, and so on) can be applied to seismographic maps, to pressure maps (whether these arise from handwriting on pressure pads or from weather data), infrared, ultraviolet, X-ray and radar images, and a variety of other cases. There is no space here to consider methods for acquisition in any of these instances and attention is concentrated on purely optical methods. In addition, space does not permit a detailed study of methods for obtaining range images using laser scanning and ranging techniques, while other methods that are specialized for 3-D work will also have to be passed by. Instead, we concentrate on (a) lighting systems for obtaining intensity images, (b) technology for receiving and digitizing intensity images, and (c) basic theory such as the Nyquist sampling theorem that underlies this type of work.

First we consider how to set up a basic system that might be suitable for the thresholding and feature detection work of Chapters 2−5.

25.2 ILLUMINATION SCHEMES

The simplest and most obvious arrangement for acquiring images is that shown in Fig. 25.1. A single source provides light over a cluster of objects on a worktable or conveyor, and this scene is viewed by a camera directly overhead. The source is typically a tungsten light that approximates to a point source. Assuming for now that the light and camera are some distance away from the objects, and are in different directions relative to them, it may be noted that:

1. different parts of the objects are lit differently, because of variations in the angle of incidence, and hence have different brightnesses as seen from the camera.
2. the brightness values also vary because of the differing absolute reflectivities[1] of the object surfaces.
3. the brightness values vary with the specularities of the surfaces in places where the incident, emergent, and phase angles are compatible with specular reflection (Chapter 15).
4. parts of the background and of various objects are in shadow and this again affects the brightness values in different regions of the image.

[1]Referring to Eq. (15.12), R_0 is the absolute surface reflectivity and R_1 is the specularity.

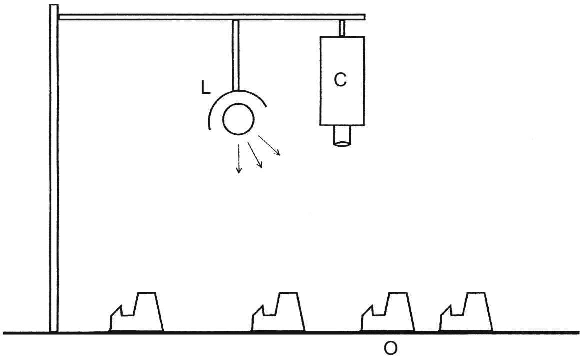

FIGURE 25.1

Simple arrangement for image acquisition: C, camera; L, light with simple reflector; O, objects on worktable or conveyor.

5. other more complex effects occur because light reflected from some objects will cast light over other objects—factors that can lead to complicated variations in brightness over the image.

Clearly, even in this apparently simple case—one point light source and one camera—the situation can become quite complex. However, effect 5 is normally reasonably marginal and is ignored in what follows. In addition, effect 3 can often be ignored except in one or two small regions of the image where sharply curved pieces of metal give rise to glints. This still leaves considerable scope for complication due to effects 1, 2, and 4.

There are two important reasons for viewing the surfaces of objects: the first is when we wish to locate objects and their facets, and the second is when we wish to scrutinize the surfaces themselves. In the first instance, it is important to try to highlight the facets by arranging that they are lit differently, so that their edges stand out clearly. In the second instance, it might be preferable to do the opposite—i.e., to arrange that the surfaces are lit very similarly, so that any variations in reflectivity caused by defects or blemishes stand out plainly. The existence of effects 1 and 2 implies that it is difficult to achieve both of these things at the same time: one set of lighting conditions is required for optimum segmentation and location, and another set for optimum surface scrutiny. In most of this book, object location has been regarded as the more difficult task and therefore the one that needs the most attention. Hence, we have imagined that the lighting scheme is set up for this purpose. In principle, a point source of light is well adapted to this situation. However, it is easy to see that if a very diffuse lighting source is employed, then angles of incidence will tend to average out and effect 2 will dominate over 1 so that, *to a first approximation*, the observed brightness values will represent variations in surface reflectance. In fact, "soft" or diffuse lighting also subdues specular reflections (effect 3), so that for the most part they can be ignored.

Returning to the case of a single point source, recall (effect 4) that shadows can become important. There is one special case when this is not so, and that is when the light is projected from exactly the same direction as the camera; we return to this case below. Shadows are a persistent cause of complications in image analysis. One problem is that it is not a trivial task to identify them, so they merely contribute to the overall complexity of any image and in particular add to the number of edges that have to be examined in order to find objects. They also make it much more difficult to use simple thresholding. (However, note that shadows can sometimes provide information that is of vital help in interpreting complex 3-D images; see, e.g., Section 15.6.)

25.2.1 Eliminating Shadows

The above considerations suggest that it would be highly convenient if shadows could be eliminated. A strategy for achieving this is to lower their contrast by using several light sources. Then the region of shadow from one source will be a region of illumination from another, and shadow contrast will be lowered dramatically. Indeed, if there are n lights, many positions of shadow will be illuminated by $n-1$ lights and their contrast will be so low that they can be eliminated by straightforward thresholding operations. However, if objects have sharp corners or concavities, there may still be small regions of shadow that are illuminated by only one light or perhaps no light at all; these regions will be immediately around the objects, and if the objects appear dark on a light background, shadows could make the objects appear enlarged or cause shadow lines immediately around them. For light objects on a dark background this is normally less of a problem.

Clearly, it seems best to aim for large numbers of lights so as to make the shadows more diffuse and less contrasting, and in the limit it appears that we are heading for the situation of soft lighting discussed earlier. However, this is not quite so. What is often required is a form of diffuse lighting that is still directional—as in the case of a diffuse source of restricted extent directly overhead: this can be provided very conveniently by a continuous ring light around the camera. This technique is found to eliminate shadows highly effectively while retaining sufficient directionality to permit a good measure of segmentation of object facets to be achieved, i.e., it is an excellent compromise although it is certainly not ideal. For these reasons, we describe its effects in some detail. In fact, it is clear that it will lead to good segmentation of facets whose boundaries lie in horizontal planes, but to poor segmentation of those whose boundaries lie in vertical planes.

The situation just described is very useful for analyzing the shape profiles of objects with cylindrical symmetry. Note, e.g., the case shown in Fig. 25.2, which involves a special type of chocolate biscuit with jam underneath the chocolate. If this is illuminated by a continuous ring light fairly high overhead (the proper

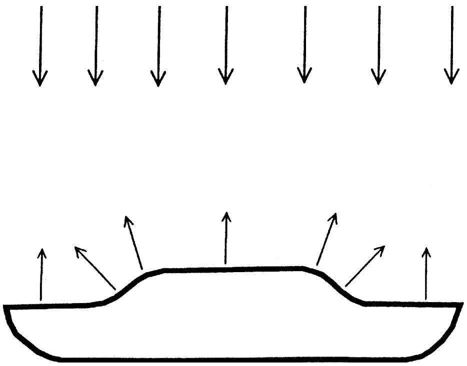

FIGURE 25.2

Illumination of a chocolate-and-jam biscuit. This figure shows the cross-section of a particular type of round chocolate biscuit with jam underneath the chocolate. The arrows show how light arriving from vertically overhead is scattered by the various parts of the biscuit.

working position), the region of chocolate above the edge of the jam reflects the light obliquely and appears darker than the remainder of the chocolate. On the contrary, if the ring light is lowered to near the worktable, the region above the edge of the jam appears *brighter* than the rest of the chocolate because it scatters light upward rather than sideways. Clearly there is also[2] a particular height at which the ring light can make the jam boundary disappear (Fig. 25.3), this height being dictated by the various angles of incidence and reflection and by the relative direction of the ring light. In comparison, if the lighting were made highly diffuse, these effects would tend to disappear and the jam boundary would always have very low contrast.

Suitable fluorescent ring lights are readily available and straightforward to use, and provide a solution that is more practicable than the alternative means of eliminating shadows mentioned above—that of illuminating objects directly from the camera direction, e.g., via a half-silvered mirror.

Earlier, the one case we did not completely solve arose when we were attempting to segment facets whose joining edges were in vertical planes. There appears to be no simple way of achieving a solution to this problem without recourse to switched lights (see Chapter 15); this option will not be discussed further here.

We have now identified various practical forms of lighting that can be used to highlight various object features and which eliminate complications as far as possible. These types of lighting are restricted in what they can achieve (as would clearly be expected from the shape-from-shading ideas of Chapter 15). However,

[2]The latter two situations are described for interest only.

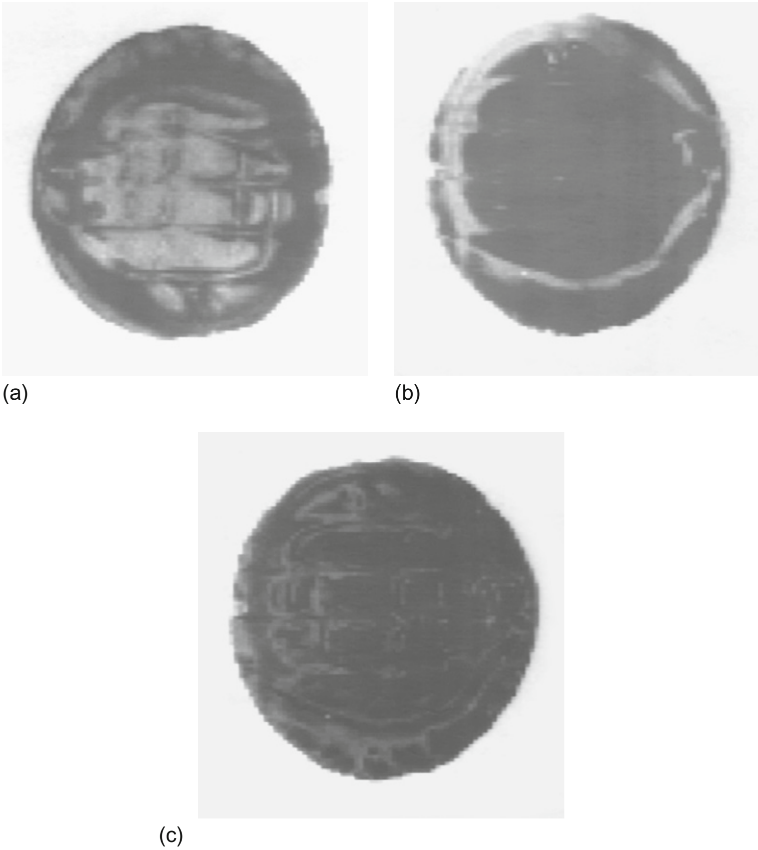

FIGURE 25.3

Appearance of the chocolate-and-jam biscuit of Fig. 25.2. (a) How the biscuit appears to a camera directly overhead when illuminated as in Fig. 25.2; (b) appearance when the lights are lowered to just above table level; (c) appearance when the lights are raised to an intermediate level making the presence of the jam scarcely detectable.

they are exceedingly useful in a variety of applications. A final problem is that two lighting schemes may have to be used in turn, the first for locating objects and the second for inspecting their surfaces. However, this problem can largely be overcome by *not* treating the latter case as a special one requiring its own lighting scheme, but rather noting the direction of lighting and *allowing for* the resulting variation in brightness values by taking account of the known shape of the object. The opposite approach is generally of little use unless other means are used for locating the object. However, the latter situation frequently arises in practice: imagine that a slab of concrete or a plate of steel is to be inspected for defects. In that case the position of the object is known and it is clearly best to set up the most uniform lighting arrangement possible, so as to be most sensitive to small variations in brightness at blemishes. This is, then, an important practical problem, to which we now turn.

25.2.2 Principles for Producing Regions of Uniform Illumination

While initially it may appear to be necessary to illuminate a worktable or conveyor uniformly, a more considered view is that a uniform flat material should *appear* uniform so that the spatial distribution of the light emanating from its surface is uniform. The relevant quantity to be controlled is therefore the radiance of the surface (light intensity in the image). Following the work of Section 15.4 relating to Lambertian (matt) surfaces, the overall reflectance R of the surface is given by:

$$R = R_0 \mathbf{s}.\mathbf{n} \tag{25.1}$$

where R_0 is the absolute reflectance of the surface and $\mathbf{n}$, $\mathbf{s}$ are unit vectors along the local normal to the surface and the direction of the light source, respectively.

Clearly, the assumption of a Lambertian surface can be questioned, since most materials will give a small degree of specular reflection, but in this section we are mainly interested in those nonshiny substances for which Eq. (25.1) is a good approximation. In any case, special provision normally has to be made for examining surfaces with a significant specular reflectance component. However, note that the continuous strip lighting systems considered below have the desirable property of very largely suppressing any specular components.

Next we recognize that illumination will normally be provided by a set of lights at a certain height h above a worktable or conveyor. We start by taking the case of a single point source at height h. Supposing that this is displaced laterally through a distance a, so that the actual distance from the source to the point of interest on the worktable is d, I will have the general form:

$$I = \frac{c \cos i}{d^2} = \frac{ch}{d^3} \tag{25.2}$$

where c is a constant factor (see Fig. 25.4).

Eq. (25.2) represents a distinctly nonuniform region of intensity over the surface. However, this problem may be tackled by providing a suitable distribution of lights. A neat solution is provided by a symmetrical arrangement of two strip lights that will clearly help to make the reflected intensity much more uniform (Fig. 25.5). We illustrate this idea by reference to the well-known arrangement of a pair of "Helmholtz" coils widely used for providing a uniform magnetic field, with the separation of the coils made equal to their radius so as to eliminate the second-order variation in field intensity.

In a similar way, the separation of the strip lights can be adjusted so that the second-order term vanishes (Fig. 25.5(b)). There is an immediate analogy also with the second-order Butterworth low-pass filter, which gives a maximally flat

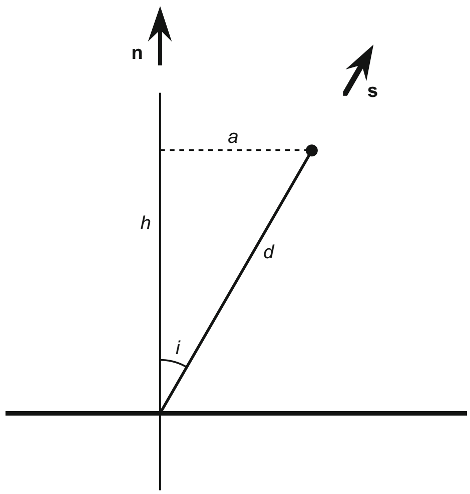

FIGURE 25.4

Geometry for a single point source illuminating a surface. Here a point light source at a height *h* above a surface illuminates a general point with angle of incidence *i*. **n** and **s** are respectively unit vectors along the local normal to the surface and the direction of the light source.

Source: © *IEE 1997*

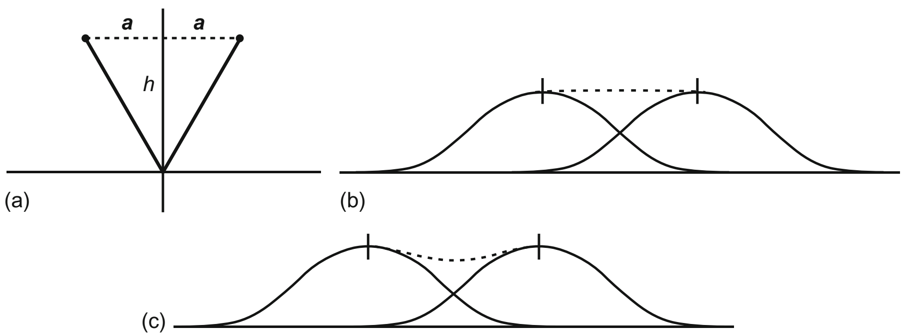

FIGURE 25.5

Effect of using two strip lights for illuminating a surface. Part (a) shows two strip lights at a height *h* above a surface, and part (b) shows the resulting intensity patterns for each of the lights; the dotted line shows the combined intensity pattern. Part (c) shows the corresponding patterns when the separation of the lights is increased slightly.

Source: © *IEE 1997*

response, the second-order term in the frequency response curve being made zero and the lowest order term then being the fourth-order term (Kuo, 1966). In fact, the latter example demonstrates how the method might be improved further—by aiming for a Chebychev type of response in which there is some ripple in the pass band, yet the *overall* pass-band response is flatter (Kuo, 1966). In a lighting application, we should aim to start with the strip lights not just far enough apart so that the second-order term vanishes, but slightly further apart so that the intensity is *almost* uniform over a rather larger region (Fig. 25.5(c)). This reflects the fact that in practice the prime aim will be to achieve a given degree of uniformity over the maximum possible size of region.

In principle it is easy to achieve a given degree of uniformity over a larger region by starting with a given response and increasing the linear dimensions of the *whole* lighting system proportionately. Although valid in principle, this approach will frequently be difficult to apply in practice: e.g., it will be limited by convenience and by availability of the strip lights; it must also be noted that as the size of the lighting system increases, so must the power of the lights. Hence, in the end we will have only one adjustable geometric parameter by which to optimize the response.

Finally, in many practical situations, it will be less useful to have a long narrow working region than one whose aspect ratio is close to unity. We shall consider two such cases—a circular ring light and a square ring light. The first of these is conveniently provided in diameters of up to at least 30 cm by commercially available fluorescent tubes, while the second can readily be constructed—if necessary on a very much larger scale—by assembling a set of four linear fluorescent tubes. In this case we take the tubes to be finite in length, and in contact at their ends: these are made into an assembly that can be raised or lowered to optimize the system. Thus, these two cases have fixed linear dimensions characterized in each case by the parameter a, and it is h that is adjusted rather than a. To make comparisons easier, we assume in all cases that a is the constant and h is the optimization parameter (Fig. 25.6).

25.2.3 Case of Two Infinite Parallel Strip Lights

First we take the case of two infinite parallel strip lights. In this case, the intensity I is given by the sum of the intensities I_1, I_2 for the two tubes:

$$I_1(x) = h \int_{-\infty}^{\infty} [(a-x)^2 + (v-y)^2 + h^2]^{-3/2} \mathrm{d}v \tag{25.3}$$

$$I_2(x) = I_1(-x) \tag{25.4}$$

Suitable substitutions permit Eq. (25.3) to be integrated, and the final result is:

$$I = \frac{2h}{(a-x)^2 + h^2} + \frac{2h}{(a+x)^2 + h^2} \tag{25.5}$$

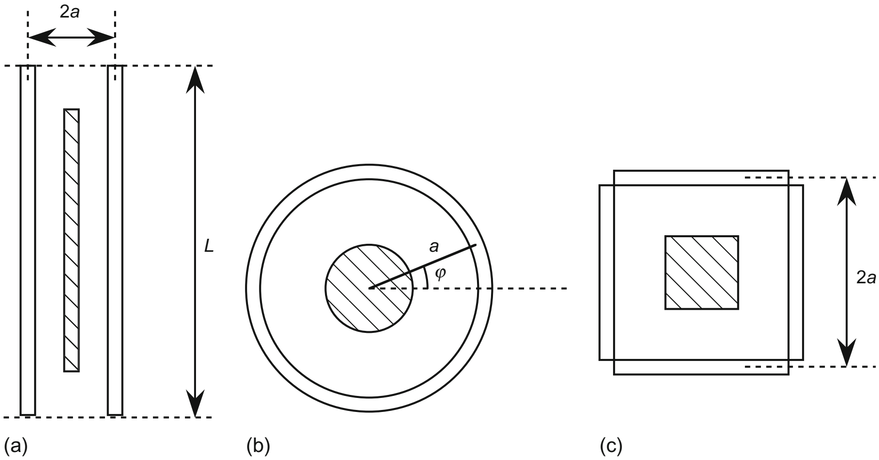

FIGURE 25.6

Lighting arrangements for obtaining uniform intensity. This diagram shows three arrangements of tubular lights for providing uniform intensity over fairly large regions, shown cross-hatched in each case. Part (a) shows two long parallel strip lights, part (b) shows a circular ring light, and part (c) shows four strip lights arranged to form a square "ring." In each case, height *h* above the worktable must also be specified.

Source: © *IEE 1997*

Differentiating I twice and setting $\mathrm{d}^2I/\mathrm{d}x^2 = 0$ at $x = 0$ eventually (Davies, 1997b) yields the maximally flat condition:

$$h = \sqrt{3}a \tag{25.6}$$

However, as noted above, it should be better to aim for minimum *overall* ripple over a region $0 \le x \le x_1$. The situation is shown in Fig. 25.7. We take the ripple ΔI as the difference in height between the maximum intensity I_m and the minimum intensity I_0, at $x = 0$, and on this basis the maximum permissible deviation in x is the value of x where the curve again crosses the minimum value I_0.

A simple calculation shows that the intensity is again equal to I_0 for $x = x_1$, where:

$$x_1 = (3a^2 - h^2)^{1/2} \tag{25.7}$$

the graph of h vs. x_1 being the circle $x_1^2 + h^2 = 3a^2$ (Fig. 25.8, top curve). Interestingly, the maximally flat condition is a special case of the new one, applying where $x_1 = 0$.

Further mathematical analysis of this case is difficult: numerical computation leads to the graphs presented in Fig. 25.8. The top curve in Fig. 25.8 has already been referred to, and shows the optimum height for selected ranges of values of x up to x_1. Taken on its own, this curve would be valueless as the accompanying nonuniformity in intensity would not be known. This information is provided by

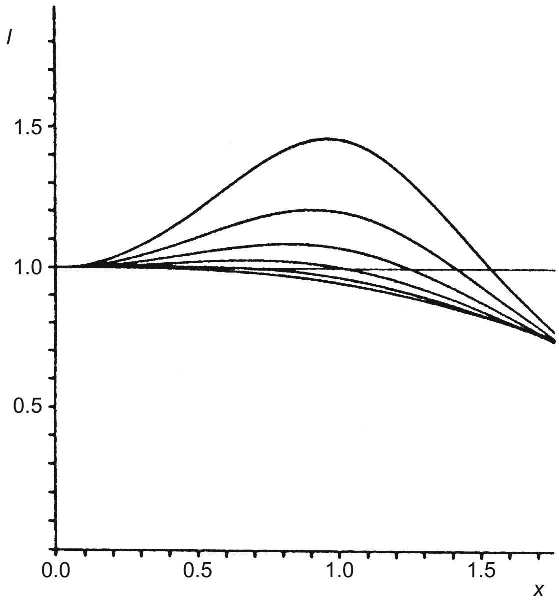

FIGURE 25.7

Intensity variation for two infinite parallel strip lights. This diagram shows the intensity variation I as a function of the distance x from the center of symmetry for six different values of h; h increases in steps of 0.2 from 0.8 for the top curve to 1.8 for the bottom curve. The value of h corresponding to the maximally flat condition is $h = 1.732$. x and h are expressed in units of a, while I is normalized to give a value of unity at $x = 0$.

Source: © *IEE 1997*

the left curve in Fig. 25.8. However, for design purposes, it is most important first to establish what range of intensities accompanies a given range of values of x, since this information (Fig. 25.8, bottom curve) will permit the necessary compromise between these variables to be made. Having decided on particular values of x_1 and ΔI, the value of the optimization parameter h can then be determined from one of the other two graphs: both are provided for convenience of reference (once two of the graphs are provided, the third gives no completely new information). Maximum acceptable variations in ΔI are assumed to be in the region of 20%, although the plotted variations are taken up to ~50% to give a more complete picture; on the other hand, in most of the practical applications envisaged here, ΔI would be expected not to exceed 2–3% if accurate measurements of products are to be made.

The ΔI vs. x_1 variation varies faster than the fourth power of x_1, there being a very sharp rise in ΔI for higher values of x_1. This means that once ΔI has been specified for the particular application, there is little to be gained by trying to

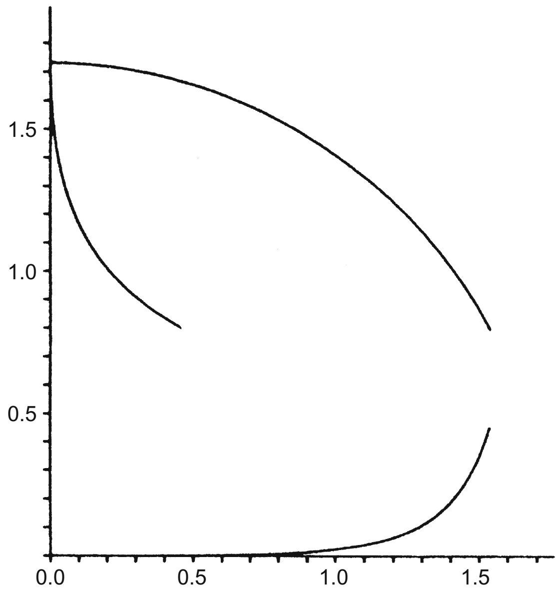

FIGURE 25.8

Design graphs for two parallel strip lights. Top, h vs. x_1. Left, h vs. ΔI. Bottom, ΔI vs. x_1. The information in these graphs has been extracted from Fig. 25.7. In design work, a suitable compromise working position would be selected on the bottom curve, and then h would be determined from one of the other two curves. In practice, ΔI is the controlling parameter, so the left and bottom curves are the important ones, the top curve containing no completely new information.

Source: © *IEE 1997*

squeeze extra functionality through going to higher values of x_1, i.e., in practice ΔI is the controlling parameter.

In the case of a circular ring light, the mathematics is more tedious (Davies, 1997b) and it is not profitable to examine it here. The final results are very similar to those for parallel strip lights. They would be used for design in the identical manner to that outlined earlier for the previous case.

In the case of a square ring light, the mathematics is again tedious (Davies, 1997b), but the results follow the same pattern and warrant no special comment.

25.2.4 Overview of the Uniform Illumination Scenario

Previous work on optical inspection systems has largely ignored the design of optimal lighting schemes. This section has tackled the problem in a particular case of interest—how to construct an optical system that makes a uniform matt surface appear uniformly bright, so that blemishes and defects can readily be detected with minimal additional computation. Three cases for which calculations have been carried out cover a good proportion of practical lighting schemes, and the design principles described here should be applicable to most other schemes that could be employed.

The results are best presented in the form of graphs. In any one case, the graph shows the tradeoff between variation in intensity and range of position on the working surface, from which a suitable working compromise can be selected. The other two graphs provide data for determining the optimization parameter (the height of the lights above the working surface).

Clearly, a wide variety of lighting arrangements are compatible with the general principles presented above: thus it is not worthwhile to give any detailed dimensional specifications here. However, the adjustment of just one parameter (the height) permits uniform illumination to be achieved over a reasonable region. Note that (as for a Chebyshev filter) it may be better to arrange a slightly less uniform brightness over a larger region than absolutely uniform brightness over a small region. It is left to empirical tests to finalize the details of the design.

Finally, it should be reiterated that such a lighting scheme is likely to be virtually useless for segmenting object facets from each other—or even for discerning relatively low curvatures on the surface of objects: its particular value lies in the scrutiny of surfaces via their absolute reflectivities, without the encumbrance of switched lights (see Chapter 15). It should also be emphasized that the aim of the discussion in the past few sections has been to achieve as much as possible with a simple static lighting scheme set up systematically. Naturally, such solutions are compromises and again no substitute for the full rigor of switched lighting schemes.

25.2.5 Use of Line-Scan Cameras

Throughout the above discussion it has been assumed implicitly that a conventional "area" camera is employed to view the objects on a worktable. However, when products are being manufactured in a factory they are very frequently moved from one stage to another on a conveyor. Stopping the conveyor to acquire an image for inspection would impose unwanted design problems: for this reason use is made of the fact that the speed of the conveyor is reasonably uniform, and an area image is built up by taking successive linear snapshots. This is achieved with a line-scan camera that consists of a row of photocells on a single integrated circuit sensor; the orientation of the line of photocells must of course be normal to the direction of motion. The internal design of line-scan and other cameras is discussed further below. However, we here concentrate on the lighting arrangement to be used with such a camera.

When using a line-scan camera, it is natural to select a lighting scheme that embodies the same symmetry as the camera: indeed, the most obvious such scheme is a pair of long fluorescent tubes parallel to the line of the camera (and perpendicular to the motion of the conveyor). We here caution against this "obvious" scheme, since a small round object, e.g., will not be lit symmetrically. Of course, there are difficulties in considering this problem in that different parts of the object are viewed by the line-scan camera at different moments, but note that for small objects a linear lighting scheme will not be isotropic: this could lead to small distortions being introduced in measurements of object dimensions. This means that in practice the ring and other symmetrical lighting schemes described above are likely to be more closely optimal even when a line-scan camera is used. For larger objects much the same situation applies, although the geometry is more complex to work out in detail.

Finally, the comment above that conveyor speeds are "reasonably uniform" should be qualified. The author has come across cases where this is true only as a first approximation. As with many mechanical systems, conveyor motion can be unreliable: e.g., it can be jerky, and in extreme cases not even purely longitudinal! Such circumstances frequently arise through a variety of problems that cause slippage relative to the driving rollers—the effects of wear or of an irregular join in the conveyor material, misalignment of the driving rollers, and so on. Furthermore, the motors controlling the rollers may not operate at constant speed, either in the short term (e.g., because of varying load) or in the longer term (e.g., because of varying mains frequency and voltage). While, therefore, it cannot be assumed that a conveyor will operate in an ideal way, careful mechanical design can minimize these problems. However, when high accuracy is required, it will be necessary to monitor the conveyor speed, perhaps by using optically coded disk devices, and feeding appropriate distance marker pulses to the controlling computer. Even with this method, it will be difficult to match in the longitudinal direction the extremely high accuracy[3] available from the line-scan camera in the lateral direction. However, images of 512×512 pixels that are within 1 pixel accuracy in each direction should normally be available.

25.2.6 Light Emitting Diode (LED) Sources

The past decade has seen a rapidly changing situation in the types of lighting available for various applications. The earlier tungsten lights coexisted for a long time with fluorescent tubes, and more recently compact fluorescents and halogen lights have moved forward. Low-power LEDs have been available for a significant time, but were initially only suitable as indicator lights rather than to provide illumination. In fact, there were problems in bringing them to higher power levels, and at the same time making their cost competitive. However, this position has been changing rapidly, and LED headlights and sidelights are now ubiquitous

[3]Line-scan cameras are available with 4096 or greater numbers of photocells in a single linear array. In addition, these arrays are fabricated using very high precision technology (see Section 25.3), so considerable reliance can be placed on the data they provide.

on road vehicles. To some extent the power problems have been solved by employing *arrays* of LEDs—a trend that has been very evident with vehicle lights. For inspection, LEDs now seem to provide the main route forward: for instance, they do not have the high-frequency firing problems of fluorescents, or the unreliability and short lives of fluorescents, tungstens and halogens. And the need for multiple LEDs to provide high illumination levels is synergistic with the need for uniform lighting, so that any illumination shape profile that would be useful for inspection, and could earlier have been provided by long or circular fluorescents or fiber-optic bundles, can now be achieved using arrangements of LEDs. While the home lighting market is still waiting for high-power LEDs to come down in price, they are easily within reach of important inspection applications. It should also be noted that LEDs are commonly guaranteed for 7 years, although their real lifetime is closer to 20 years or more (this figure assumes up to $\sim$50% usage per day). These long lifetimes are offset to some extent by gradually falling emission (a drop of 10% in $\sim$20,000 hours),[4] but this can be cancelled out by slightly derating and progressively increasing the dc supply current (the light intensity emitted by LEDs is directly proportional to the supply current). Overall, the relevant advantages of LEDs are easily controllable intensity, directly proportional to current; output that does not switch on and off at a rate determined by the mains frequency; exceptionally long life; high conversion efficiency; and an intense light from a small area that is readily focussed.

25.3 CAMERAS AND DIGITIZATION

For a good many years the camera that was normally used for image acquisition was the TV camera with a vidicon or related type of vacuum tube. The scanning arrangements of such cameras became standardized, first to 405 lines, then to 625 lines (or 525 lines in the United States). In addition, it is usual to interlace the image—i.e., to scan odd lines in one frame and even lines in the next frame, then repeat the process, each full scan taking 1/25 second (1/30 second in the United States). There are also standardized means for synchronizing cameras and monitors, using line and frame "sync" pulses. Thus, the vacuum TV camera left a legacy of scanning techniques that are in the process of being eliminated with the advent of digital TV.[5] However, as the result of the legacy is still present, we include a few more details here.

It is important to note that the output of these early cameras is inherently analog, consisting of a continuous voltage variation, although this applies only along

[4]A useful view of the situation appears on the following manufacturer's (Philips Colour Kinetics') website: http://www.colorkinetics.com/support/whitepapers/LEDLifetime.pdf (website accessed 23 July 2011).
[5]All the vestiges of the old system will not have been swept away until all TV receivers and monitors as well as the cameras are digital.

the line directions: the scanning action is discrete in that lines are used, making the output of the camera part analog and part digital. Hence, before the image is available as a set of discrete pixels, the analog waveform has to be sampled. Since some of the line scanning time is taken up with frame synchronization pulses, only about 550 lines are available for actual picture content. In addition, the aspect ratio of a standard TV image is 4:3 and it is common to digitize TV pictures as 512×768 or 512×512 pixels. Note also that after the analog waveform has been sampled and pixel intensity values have been established, it is still necessary to digitize the intensity values.

Modern solid-state cameras are much more compact and robust, and generate less noise; a very important additional advantage is that they are not susceptible to distortion,[6] because the pixel pattern is fabricated very accurately by standard integrated circuit photolithography techniques. They have thus replaced vacuum tube cameras in all except special situations.

Most solid-state cameras currently available are of the self-scanned charge-coupled device (CCD) type and attention is concentrated on these in what follows. In a solid-state CCD camera, the target is a piece of silicon semiconductor that possesses an array of photocells at the pixel positions. Hence, this type of camera digitizes the image from the outset, although in one respect—that signal amplitude represents light intensity—the image is still analog. The analog voltages (or more accurately the charges) representing the intensities are systematically passed along analog shift registers to the output of the instrument, where they may be digitized to give 6−8 bits of grayscale information (the main limitation here being lack of uniformity among the photosensors rather than noise *per se*). This architecture is important as it means that widely available CCD cameras can be triggered and read out at any desired rate by externally applied pulses.

Interestingly, the old vacuum tube TV cameras had a spectral response curve that peaked at much the same position as the spectral pattern of ("daylight") fluorescent lights—which itself matches the response of the human eye (Table 25.1). However, CCD cameras have significantly lower response to the spectral pattern of fluorescent tubes (Table 25.1). In general this may not matter too greatly, but when objects are moving, the integration time of the camera is limited and sensitivity can suffer. In such cases the spectral response is an important factor and may dictate against use of fluorescent lights (this is particularly relevant where CCD line-scan cameras are used with fast-moving conveyors).

An important factor in the choice of cameras is the delay lag that occurs before a signal disappears. This clearly causes problems with moving images. Fortunately, the effect is entirely eliminated with CCD camera, since the action of reading an image wipes the old image. However, moving images require frequent reading and this implies loss of integration time and therefore loss of sensitivity—a factor that normally has to be made up by increasing the power of

[6]However, this does not prevent distortions from being introduced by other mechanisms—poor optics, poor lighting arrangements, perspective effects, and so on.

Table 25.1 Spectral Responses

Device	Band (nm)	Peak (nm)
Vidicon	200–800	~550
CCD	400–1000	~800
Fluorescent tube	400–700	~600
Human eye	400–700	~550

In this table, the response of the human eye is included for reference. Note that the CCD response peaks at much higher wavelength than the vidicon or fluorescent tube, and therefore is often at a disadvantage when used in conjunction with the latter.

illuminating sources. Camera "burn-in" is another effect that is absent with CCD cameras but which causes severe problems with certain types of conventional camera: it is the long-term retention of picture highlights in the light-sensitive material, which makes it necessary to protect the camera against bright lights and to take care to make use of lens covers whenever possible. Finally, "blooming" is the continued generation of electron-hole pairs even when the light-sensitive material is locally saturated with carriers, with the result that the charge spreads and causes highlights to envelop adjacent regions of the image. Both CCD and conventional camera tubes are subject to this problem, although it is inherently worse for CCDs, and this has led to the production of antiblooming structures in these devices. Space precludes detailed discussion of the situation here.

25.3.1 Digitization

The remaining important item to be studied in this context is that of digitization, i.e., conversion of the original analog signals into digital form. There are many types of analog-to-digital converter (ADC) but the ones that are used for digitizing images have so much data to process—usually in a very short time if real-time analysis is called for—that special types have to be employed. The only one considered here is the "flash" ADC, so called because it digitizes all bits simultaneously, in a flash. In fact, it possesses $n - 1$ analog comparators to separate n gray levels, followed by a priority encoder to convert the $n - 1$ results into normal binary code. Such devices produce a result in a very few nanoseconds and their specifications are generally quoted in megasamples per second (typically in the range 50–200 megasamples/second). For some years these were available only in 6-bit versions (apart from some very expensive parts), but nowadays 8-bit versions are available at extremely low cost:[7] such 8-bit devices are probably sufficient for most needs considering that a certain amount of sensor noise, or

[7]Indeed, as is clear from the advent of cheap web cameras and digital cameras, it is becoming virtually impossible to get noncolor versions of such devices.

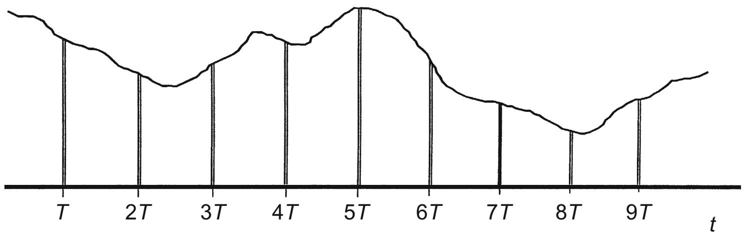

FIGURE 25.9

The process of sampling a time-varying signal: a continuous time-varying 1-D signal is sampled by narrow sampling pulses at a regular rate $f_r = 1/T$, which must be at least twice the bandwidth of the signal.

variability, is usually present below these levels and that it is difficult to engineer lighting to this accuracy. In fact, devices with even greater grayscale resolution can be obtained: these are useful for extending the overall dynamic range capability that is required when lighting levels are highly variable (which is the normal occurrence outdoors during the course of the day).

25.4 THE SAMPLING THEOREM

The Nyquist sampling theorem underlies all situations where continuous signals are sampled and is especially important where patterns are to be digitized and analyzed by computers. This makes it highly relevant both with visual patterns and with acoustic waveforms, hence it is described briefly in this section.

Consider the sampling theorem first in respect of a 1-D time-varying waveform. The theorem states that a sequence of samples (Fig. 25.9) of such a waveform contains all the original information and can be used to regenerate the original waveform exactly, but only if (a) the bandwidth W of the original waveform is restricted and (b) the rate of sampling f is at least twice the bandwidth of the original waveform—i.e., $f \geq 2W$. Assuming that samples are taken every T seconds, this means that $1/T \geq 2W$.

At first it may be somewhat surprising that the original waveform can be reconstructed exactly from a set of discrete samples. However, the two conditions for achieving this are very stringent. What they are demanding in effect is that the signal must not be permitted to change unpredictably (i.e., at too fast a rate), else accurate interpolation between the samples will not prove possible (the errors that arise from this source are called "aliasing" errors).

Unfortunately, the first condition is virtually unrealizable, since it is close to impossible to devise a low-pass filter with a perfect cut-off. Recall from Chapter 3 that a low-pass filter with a perfect cut-off will have infinite extent in the time domain, so any attempt at achieving the same effect by time domain operations must be doomed to failure. However, acceptable approximations can

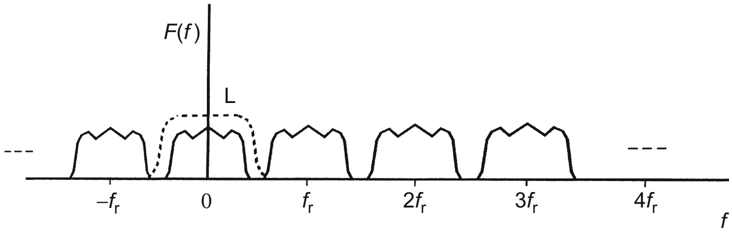

FIGURE 25.10

Effect of low-pass filtering to eliminate repeated spectra in the frequency domain (f_r, sampling rate; L, low-pass filter characteristic). This diagram shows the repeated spectra of the frequency transform $F(f)$ of the original sampled waveform. It also demonstrates how a low-pass filter can be expected to eliminate the repeated spectra to recover the original waveform.

be achieved by allowing a "guard band" between the desired and actual cut-off frequencies. This means that the sampling rate must be higher than the Nyquist rate (in telecommunications, satisfactory operation can generally be achieved at sampling rates around 20% above the Nyquist rate—see Brown and Glazier, 1974).

One way of recovering the original waveform is by applying a low-pass filter. This approach is intuitively correct, since it acts in such a way as to broaden the narrow discrete samples until they coalesce and sum to give a continuous waveform. Indeed, this method acts in such a way as to eliminate the "repeated" spectra in the transform of the original sampled waveform (Fig. 25.10). This in itself shows why the original waveform has to be narrow-banded before sampling, so that the repeated and basic spectra of the waveform do not cross over each other and become impossible to separate with a low-pass filter. The idea may be taken further because the Fourier transform of a square cut-off filter is the sinc ($\sin u/u$) function (Fig. 25.11). Hence, the original waveform may be recovered by convolving the samples with the sinc function (which in this case means replacing them by sinc functions of corresponding amplitudes). This has the effect of broadening out the samples as required, until the original waveform is recovered.

So far we have considered the situation only for 1-D time-varying signals. However, recalling that there is an exact mathematical correspondence between time and frequency domain signals on the one hand and spatial and spatial frequency signals on the other, the above ideas may all be applied immediately to each dimension of an image (although the condition for accurate sampling now becomes $1/X \geq 2W_X$, where X is the spatial sampling period and W_X is the spatial bandwidth). Here we accept this correspondence without further discussion and proceed to apply the sampling theorem to image acquisition.

Consider next how the signal from a camera may be sampled rigorously according to the sampling theorem. First, note that this has to be achieved both horizontally and vertically. Perhaps the most obvious solution to this problem is to perform the process optically, perhaps by defocusing the lens; however, the optical

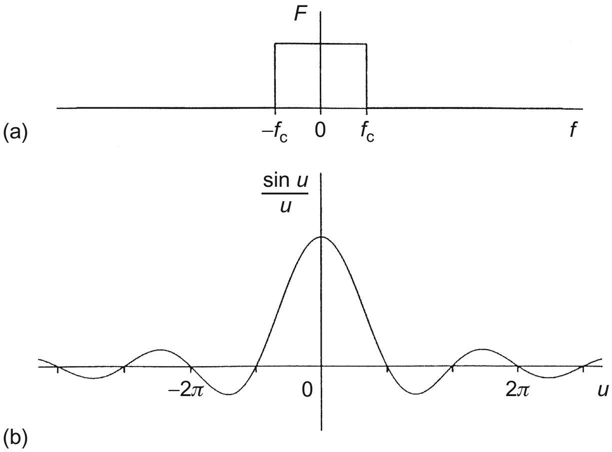

FIGURE 25.11

The sinc (sin u/u) function shown in (b) is the Fourier transform of a square pulse (a) corresponding to an ideal low-pass filter. In this case, $u = 2\pi f_c t$, f_c being the cut-off frequency.

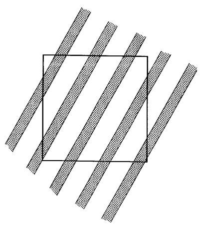

FIGURE 25.12

Low-pass filtering carried out by averaging over the pixel region: an image with local high-frequency banding is to be averaged over the whole pixel region by the action of the sensing device.

transform function for this case is frequently (i.e., for extreme cases of defocusing) very odd, going negative for some spatial frequencies and causing contrast reversals; hence, this solution is far from ideal (Pratt, 2001). Alternatively, we could use a diffraction-limited optical system or perhaps pass the focussed beam through some sort of patterned or frosted glass to reduce the spatial bandwidth artificially. None of these techniques will be particularly easy to apply, nor (apart possibly from the second) will it give accurate solutions. However, this problem is not as serious as might be imagined. If the sensing region of the camera (per pixel) is reasonably large, and close to the size of a pixel, then the averaging inherent in obtaining the pixel intensities will in fact perform the necessary narrow-banding (Fig. 25.12). To analyze the situation in more detail, note that a pixel is essentially square with a sharp cut-off at its borders. Thus its spatial frequency pattern is a 2-D sinc function, which (taking the central positive peak) approximates to a low-pass spatial frequency filter. This approximation improves somewhat as the border between pixels becomes more fuzzy.

The point here is that the worst case from the point of view of the sampling theorem is that of extremely narrow discrete samples, but clearly this worst case is most unlikely to occur with most cameras. However, this does not mean that sampling is automatically ideal—and indeed it is not, since the spatial frequency pattern for a sharply defined pixel shape has (in principle) infinite extent in the spatial frequency domain. The review by Pratt (2001) clarifies the situation and shows that there is a tradeoff between aliasing and resolution error. Overall, quality of sampling will be one of the limiting factors if the greatest precision in image measurement is aimed for: if the bandwidth of the presampling filter is too low, resolution will be lost; if it is too high, aliasing distortions will creep in; and if its spatial frequency response curve is not suitably smooth, a guard band will have to be included and performance will again suffer.

25.5 HYPERSPECTRAL IMAGING

For a number of decades, multispectral imaging has been employed in remote sensing to obtain sufficient data to separate various land regions, such as soil, water, crops, roads, and buildings. In particular, multispectral data typically consists of four to six channels of color; infrared and microwave data from which relevant regions can be separated, e.g., with the aid of principal components analysis (PCA). However, for some time there has been a need in certain applications for even more image data, and this has led to the development of so-called "hyperspectral imaging." This is a generalization of multispectral data to cover a far larger number of spectral channels. In fact, hyperspectral imaging collects a whole *spectrum* of data for each pixel by application of suitable spectrometers. As a result, a "hyperspectral cube" of image data is formed. In contrast with multispectral images which contain at most tens of spectral bands at relatively isolated wavelengths, hyperspectral images contain hundreds or even thousands of contiguous channels. This is useful as it permits detailed selections of the required data to be made some time after the images have been obtained. Clearly, the technique is costly in storage and in the processing load needed both to select appropriate subsets of the data and to perform subsequent processing: the fundamental problem is that typical hyperspectral images usually contain hundreds of megabytes of data. Nevertheless, the technique is well adapted to remote sensing of crops from satellites, and as often happens, once a new technology develops, workers see the opportunity to apply it to their own particular types of data, and the technology is forced to develop further, for reliability, ease of use, and reduction of cost. As in the case of MRI, which is applied mainly but by no means exclusively to medical imaging, this new technology is poised to be applied extremely widely. Already it is being tested and applied to the inspection of fruit and vegetables, and to the sensitive detection of nematode worms in fish (Gómez et al., 2007; Heia et al., 2007).

In such applications, the objects under inspection are moved slowly, as would be normal for line-scan camera inspection, but here each line of pixels is scanned over the whole spectrum to image all the color channels. Thus, instead of a line of pixels being grabbed, a whole plane of pixels is grabbed in which the additional coordinate represents the color. This can be achieved in two ways: (1) the same line-scan camera is used and the spectrum is scanned sequentially over it using a spectrometer with a rotating mirror or rotating diffraction grating; (2) an area camera is used to view the whole spectral plane in parallel (this could in principle be carried out with a prism and an area camera). Clearly, use of standard area cameras applied sequentially in ~10 nm steps over the required spectral range is also a possibility. Typical applications employ visible and infrared measurements over a range 500−1000 nm, or infrared measurements over the range 960−1700 nm using an InGaAs NIR camera (Qin and Lu, 2008; Mahesh et al., 2008). Further details will not be considered here, because of the particularly rapid development of the subject.

25.6 CONCLUDING REMARKS

This chapter has aimed to give some background to the problems of acquiring images, particularly for inspection applications. Methods of illumination have been deemed to be worthy of considerable attention since they furnish means by which the practitioner can help to ensure that an inspection system operates successfully—and indeed that its vision algorithms are not unnecessarily complex, thereby necessitating excessive hardware expense for real-time implementation. Means of arranging reasonably uniform illumination and freedom from shadows have been taken to be of significant relevance and allotted fair attention (interestingly, these topics are scarcely mentioned in most books on this subject—a surprising fact that perhaps indicates that few authors ascribe much importance to this vital aspect of the work). For recent publications on illumination and shadow elimination, see the following section.

By contrast, camera systems and digitization techniques have been taken to be purely technical matters to which little space could be devoted (to be really useful—and considering that most workers in this area buy commercial cameras and associated frame-grabbing devices ready to plug into a variety of computers—whole chapters would have been required for each of these topics). Because of its theoretical importance, it was relevant to give some background to the sampling theorem and its implications although, considering the applications covered in this book, only limited space could be devoted to this topic (see Rosie, 1966 and Pratt, 2001 for further details).

Finally, the chapter included a section on hyperspectral imaging, which has recently been the subject of much research and development. In fact, the motivation for hyperspectral imaging, with the large amount of data it can bring to bear in fields such as agriculture and inspection, has only been made thinkable as a result of the vastly increased storage and processing capabilities of modern computers. There seems little doubt that progress in its technology and application will accelerate markedly in the next few years.

> Computer vision systems are commonly highly dependent on the quality of the incoming images. This chapter has shown that image acquisition can often be improved, particularly for inspection applications, by arranging regions of nearly uniform illumination so that shadows and glints are suppressed, and vision algorithms can be simplified and speeded up.

25.7 BIBLIOGRAPHICAL AND HISTORICAL NOTES

It is a regrettable fact that very few papers and books give details of lighting schemes that are used for image acquisition, and even fewer give any rationale or background theory for such schemes. Hence, some of the present chapters appear to have broken new ground in this area. However, Batchelor et al. (1985) and Browne and Norton-Wayne (1986) give much useful information on light sources, filters, lenses, light guides, and so on, thereby complementing the work of this chapter (indeed, the former of these books gives a wealth of detail on how unusual inspection tasks, such as those involving internal threads, may be carried out).

Details of various types of scanning system, camera tube, and solid-state (e.g., CCD) device are widely available—see, e.g., Batchelor et al. (1985), and also various manufacturers' catalogs. Note that much of the existing CCD imaging device technology dates from the mid-1970s (Barbe, 1975; Weimer, 1975) and is still undergoing development.

The sampling theorem is well covered in very many books on signal processing (see, e.g., Rosie, 1966). However, details of how band-limiting should be carried out prior to sampling are not so readily available. Only a brief treatment is given in Section 25.4; for further details the reader is referred to Pratt (2001) and references contained therein.

The work described in Section 25.2 arose from the author's work on food product inspection, which required carefully controlled lighting to facilitate measurement, improve accuracy, and simplify (and thereby speed up) the inspection algorithms (Davies, 1997b). Similar motivation drove Yoon et al. (2002) to attempt to remove shadows by switching different lights on and off, and then using logic to eliminate the shadows: under the right conditions, it was only necessary to find the maximum of the individual pixel intensities between the various images. However, in outdoor scenes, it is difficult to control the lighting: instead various rules must be worked out for minimizing the problems. Prati et al. (2001) have made a comparative evaluation of available methods. Other work on shadow location and elimination has been reported by Rosin and Ellis (1995), Mikić et al. (2000), and Cucchiara et al. (2003).

In addition, Koch et al. (2001) have presented results on the use of switched lights to maintain image intensity irrespective of changes of ambient illumination, and to limit the overall dynamic range of image intensities so that the risk of over- or underexposing the scene is drastically reduced. See also the bibliography in Section 15.13 for related solutions under the guise of photometric stereo.

25.7.1 More Recent Developments

Hyperspectral imaging has developed almost explosively since 2000. Not only have the techniques themselves developed in phase with the increasing power of the PC, but also significant numbers of applications have started to come forward. In particular, Gómez et al. (2007) and Gómez-Sanchis et al. (2008) have applied it to citrus fruit inspection; Qin and Lu (2008) have applied it more generally to fruit and vegetable inspection; Heia et al. (2007) have applied it to the detection of nematode worms in cod fillets (the parasites sometimes being detected well below the surface of the fillet by this means); and Mahesh et al. (2008) have applied it to differentiate Canadian wheat classes. Outside the food inspection arena, Yuen and Richardson (2010) have developed it for security, surveillance, and target acquisition. Here the potentialities are enormous, as can be seen from the fact that the method even provides the capability for assessing human stress by monitoring blood hemoglobin oxygenation levels in the face region. For further details, see Section 25.5 and the original papers mentioned above. Note that Yeen and Richardson's paper provides interesting details of the instrumentation and instrumentation schemes needed for hyperspectral imaging hardware, which range from dispersive spectrographs to narrow-band tunable filters.