Rosenfeld, A. and Kak, A. 1982 *Digital Picture Processing*, Academic Press, New York.
Sanderson, R.J. 1983 *Machine Vision Systems: A Summary and Forecast*, Tech Tran Corp. Illinois.
Spoehr, K.T. and Lehmkuhle, S.W. 1982 *Visual Information Processing*, Freeman and Co., San Francisco.
Suresh, B., Fundakowski, R.A., Levitt T.S. and Overland, J.E. 1983 'A real-time automated visual inspection system for hot steel slabs', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-5, No. 6, pp. 563–72.
Vernon, D. (Rapporteur) 1987 *Sensoric Feedback for Control and Decision Support Systems*, Report of a Workshop in the Framework of ESPRIT Conference 1987, Directorate General XIII (eds.).
Vernon, D. 1989 'Computers See the Light', *Technology Ireland*, Vol. 21, No. 1, pp. 21–3.

2

Illumination and sensors

2.1 *Illumination*

Scene and object illumination play a key role in the machine vision process. The central purpose of imposing controlled constant illumination is to enhance visually the parts to be imaged so that their flaws, defects, and features are highlighted and so that their identification and classification by the vision system becomes somewhat easier. Although the choice of lighting will typically be application-dependent, some general points may be pertinent.

The common incandescent bulb is probably the simplest source of light. It is cost-effective and it is easily adjusted for light intensity; however, it generally provides directional illumination since it is, to an approximation, a point source of light. Hence, incandescent bulbs cast strong shadows which invariably cause problems for machine vision software. Special bulbs are normally required as degradation in emitted light intensity is common with age. Furthermore, incandescent bulbs emit considerable infra-red radiation; this does not cause problems for humans as we are not sensitive to such light but some camera sensors, particularly so-called CCD cameras, are sensitive and visual data can be washed out by the reflected infra-red rays.

For most machine vision applications, a diffuse source of light is the most suitable. Diffuse lighting is non-directional and produces a minimum amount of shadow. Fluorescent lighting is the simplest and most common method of obtaining diffuse illumination and is especially good for providing illumination of large areas.

In situations where the only features that need to be inspected are evident from the silhouette of the object, back-lighting is the most appropriate. Back-lighting, e.g. in the form of a light table, provides high contrast between the object and the background upon which the object rests. Its advantage is that it facilitates very simple object isolation or *segmentation*, a topic to which we will return in Chapter 5.

In some manufacturing environments, it is necessary to inspect moving objects. Depending on the characteristics of image sensor, it may be necessary to 'freeze' the motion of the object for an instant by the use of a strobe light or electronic flash. The lamp emits a short (1 ms) burst of light, thus the moving object is illuminated for a very short period and appears stationary. The activation of the strobe must be synchronized with the acquisition of the image. Alternatively, you can exploit cameras with a very fast 'shutter speed' or, rather, the electronic equivalent, a very short exposure time. The exposure is usually referred to as the integration time since it is the period over which the sensor integrates or averages the incident light. One would normally choose the latter option of a short integration time, since it is more ergonomic and less disruptive for humans.

☐ *Control of illumination and light levels*

As many inspection systems base much of their analysis on the absolute intensity of the incident light, the control of the object illumination can be important. In particular, if image processing and analysis decisions are being made on the basis of a fixed intensity datum (threshold), then some problems will occur if the illumination, and hence the reflected light, changes. If possible, the vision system should be able to adapt to such changes, although this does not necessarily mean that it should be capable of dealing with dynamic changes. Most illumination systems degrade quite slowly over time and it would be quite satisfactory if the system were capable of self-calibration at the beginning of each day.

Other alternatives exist, however, to this adaptive approach. One solution is to ensure that the illumination does, in fact, remain constant by monitoring it using light meters and adjusting the illumination system appropriately. Alternatively, the aperture of the camera lens might be altered. Note, however, that electronically controlled aperture lenses (so-called auto-iris lenses) should not be employed directly; their function is to alter aperture so that the average amount of light passing through the lens remains constant. This is not appropriate for machine vision systems as the grey-tone shade of a particular feature would vary, depending on the intensity of the ambient lighting.

It was mentioned above that incandescent lighting is not suitable for some cameras and, in general, one should ensure that the lighting system is compatible with the image sensor. For example, mains-powered lighting is inherently 'flickery' due to the a.c. characteristics of the mains electricity supply. Humans do not notice this, in general, because they effectively 'integrate' (or average) the incident illumination over a short period of time. This process also accounts for our ability to view moving objects in cinema films and perceive the motion to be continuous. Machine vision sensors, as we shall see in the next section, do not integrate in quite the same way and, when they acquire the image, the flicker can become apparent. The use of an appropriate (d.c., say) power supply can alleviate this problem, when it does occur.

2.2 Sensors

The task of a camera system is to convert an optical picture, typically representing a two-dimensional or three-dimensional scene, into some form suitable for use with electrical systems. Since the camera represents the direct link between the environment being examined and the information-processing system, it plays a particularly significant role and merits a good deal of attention.

2.2.1 Image formation: elementary optics

Before commencing a discussion of cameras, image sensors, and image acquisition components proper, we will first address a few fundamental issues on the optics of a vision system and of lenses in particular.

Lenses are required to focus part of the visual environment onto the image sensor. It is possible to construct an imaging system without lenses, using, for example, a collimated light source to project a shadow of a (small) object onto the sensor but such a configuration is not typical of the requirements of a vision system.

Lenses are defined by their *focal length* (quoted in millimetres: mm) and their *aperture* (the *f number*). These parameters determine the performance of the lens in terms of light-gathering power and magnification, and it often has a bearing on its physical size.

The focal length of a lens is a guide to the magnification it effects and its field of view. Selecting the focal length which is appropriate to a particular application is simply a matter of applying the basic lens equation:

$$\frac{1}{u}+\frac{1}{v}=\frac{1}{f}$$

where

v is the distance from the lens to the image,
u is the distance from the lens to the object,
f is the focal length.

Noting the magnification factor M is:

$$M=\frac{\text{image size}}{\text{object size}}$$

and, equivalently:

$$M=\frac{\text{image distance}}{\text{object distance}}$$

Thus:

$$f=\frac{uM}{M+1}$$

Hence, if we know the required magnification factor and the distance from the object to the lens, we can compute the required focal length.

For example, if a 10 cm wide object is to be imaged on a common 8.8 × 6.6 mm sensor from a distance of 0.5 m this implies a magnification factor of:

$$M = \frac{8.8}{100} = 0.088$$

So:

$$f = \frac{500 \times 0.088}{1.088} = 40.44$$

Thus, we require a 40.44 mm focal length lens. Typically, one would use a slightly shorter focal length (e.g. 35 mm) and accept the slight loss in resolution due to the larger field of view.

The *minimum focal distance* is the minimum distance between the front of the lens and the object at which the object can still be in focus. Generally speaking, lenses with short focal lengths have smaller minimum focal distances.

If the lens is required to focus on a relatively small object at short distances, the minimum focal distance (typically 300 mm) may be too large. In that case, an extension tube can be used to increase the distance, v, of the lens from the sensor and hence decrease the distance, u, from the lens to the object. For a given focal length, f, the lens equation stipulates that u decreases as v increases, if the image is to remain in focus.

The *f-number* is a measure of the amount of light allowed to pass through the lens and is a normalized measure of lens aperture. It is defined by the focal length divided by the diameter of the aperture. The standard scale is 1.4, 2, 2.8, 4, 5.6, 8, 11, 16; each increase *reduces* the amount of light passing through the lens by one-half.

The *depth of field* is the distance between the nearest and the furthest points of a scene which remains in acceptable focus at any one aperture setting. In general, the depth of field gets larger as the focused point moves away from the lens (given constant aperture and focal length). Also, the depth of field increases significantly as the aperture closes (i.e. as the f number increases) for any given focusing distance.

Lenses have many types of standard mounts, for example the Pentax, Olympus, Nikon bayonet mounts, but the standard on television, and CCTV (Closed Circuit TV), cameras is a screw mount called the C mount. Since there is a vast choice of 35 mm photographic lenses, it makes sense to be able to use these photographic lenses with C mount cameras and there are several types of bayonet adaptors available for C mount cameras. However, it should be remembered that these lenses are usually much more bulky than the miniature lenses which are specifically designed for CCD cameras. The reason for this is that a photographic lens is designed to image a 35 mm format and, for a given focal length, the optical surface must be much larger. CCD sensors (as we will see in the next section) are typically less than 10 mm in width and hence the optical surface can be much smaller.

Many of the visual problems which cause difficulty in interpreting a scene can often be solved by the use of simple filters on the camera lens. Filters are frequently used by the television and video industries to produce special effects, for example, the star-light filter used to make candlelight seem soft and star-like. Filters are just as often used to reduce, or remove, such effects; polaroid sunglasses are probably the most widely used 'filters' and are used to reduce glare. In machine vision, one of the most annoying and potentially disruptive problems is that due to specular reflection (mirror-like reflection on shiny objects). The use of a simple polarizing filter on the camera lens can often reduce the effect of these reflections.

Some sensors are sensitive to segments of the electromagnetic spectrum which do not convey visual data, e.g. infra-red radiation. The use of a simple infra-red blocking filter can solve this problem in a simple and effective manner.

2.2.2 Camera sensors

There are essentially two types of video camera available: one, the vidicon, is based on vacuum-tube technology and the other is based on semi-conductor technology.

Briefly, a vidicon tube is a photoconductive device which employs a photosensitive sensor layer consisting of several million mosaic cells insulated from one another on a transparent metal film (refer to the Figure 2.1). Each cell represents a small capacitor whose charge is a function of incident light. The sensor layer is scanned in a raster format with an electron beam over 625 lines in accordance with the television standard (discussed in the next section). This beam is deflected magnetically by a set of coils outside the tube bulb. The electron beam makes up charge lost through the incidence of light in individual mosaic cells and

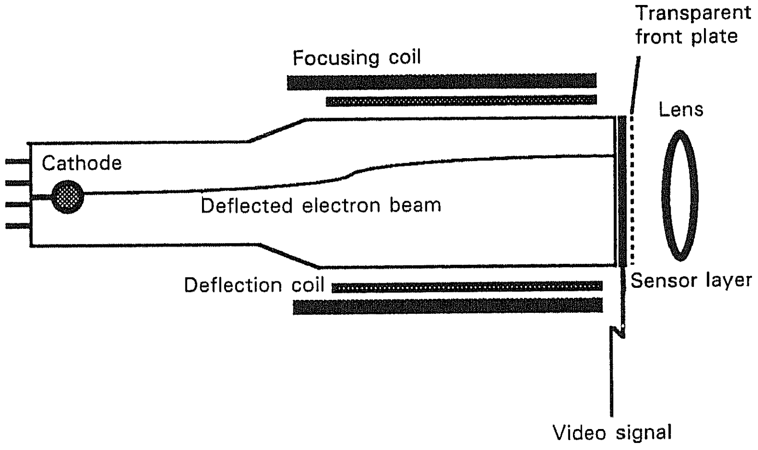

Figure 2.1 Schematic of the pickup tube of the Vidicon television camera.

so generates the video signal at the sensor element. This video signal is simply a continuous analogue signal proportional to the light intensity of the focused image. The camera electronics insert synchronization pulses (syncs.) to indicate scan lines, fields and frame ends (see Section 2.2.3).

A distinction is made between the following camera types depending on the sensor element: the standard vidicon has a sensor element comprised of antimony sulphide (Sb_2S_3), the silicon diode vidicon has a sensor element made from silicon (Si), while the plumbicon has a sensor element made of lead oxide (PbO).

Most solid-state cameras are based on charge-coupled device (CCD) technology, though there are several variations on the theme. In order for the reader to be at least acquainted with the names of these devices, they are listed here:

● charge transfer devices	CTD
● single transfer devices	STD
● bucket brigade devices	BBD
● charge coupled devices	CCD
● charge injection devices	CID
● surface charge coupled devices	SCCD
● bulk charge coupled devices	BCCD

CCD technology is the most widespread, the basic structure of which is that of an analogue shift register consisting of a series of closely spaced capacitors. Charge integration (accumulation) by the capacitors, photosites, caused by the photons comprising the incident light, provides the analogue representation of light intensity. At the end of the integration period (exposure time) these charges are read out of the sensor.

CCD sensors most commonly use one of three addressing strategies: interline transfer, frame transfer, and column–row transfer.

The interline transfer CCD is organized into column pairs of devices. An imaging column of photosensors is adjacent to an opaque vertical shift register (see Figure 2.2). Charge accumulates in the imaging column until the end of the integration period, when it is transferred to the opaque column. The signal then shifts vertically into a horizontal shift register that represents the picture sequentially, line by line. The advantage of the interline transfer is that the transfer time (to opaque storage) is short compared to the integration period. This is desirable because when transfer time approaches the integration time, solid-state sensors tend to exhibit a locally contained spreading of the image response, called smear. Thus, the interline transfer minimizes smear.

In the frame transfer organization (refer to Figure 2.3) the sensor consists of vertical columns of CCD shift registers divided into two zones. One zone, where charge accumulates during integration time, is photosensitive. When integration is complete, the whole array is transferred in parallel to the opaque storage area of the second zone.

A third type of solid-state sensor employs x–y addressing to transfer charge from the photosite to the output signal amplifier. The sensor elements are addressed

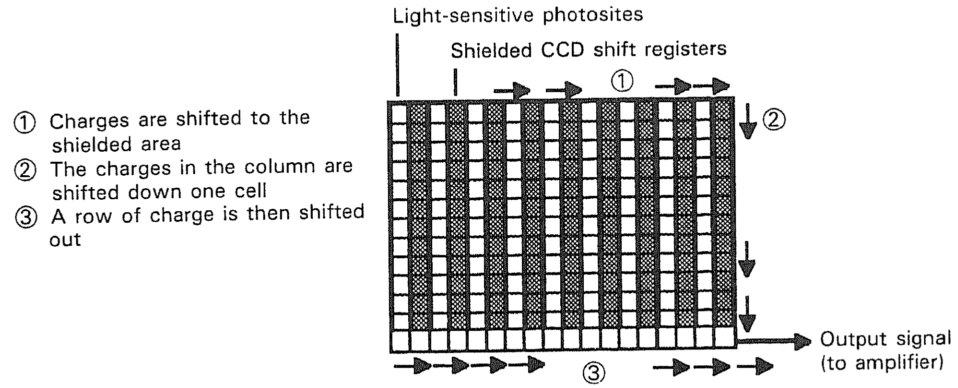

Figure 2.2 Interline transfer of charge in CCD sensors.

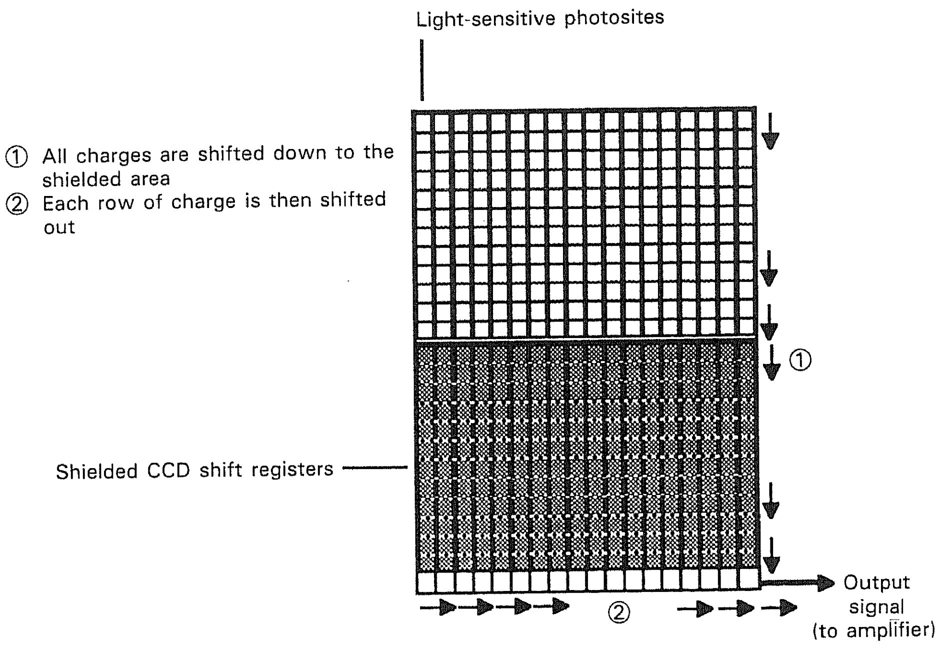

Figure 2.3 Frame transfer of charge in CCD sensors.

by selecting individual column and row electrodes. Charge collected under the column electrode is transferred to the row electrode and amplified for output.

All the devices discussed so far have been so-called 'area sensors', i.e. the sensor has been a two-dimensional array of photosites. However, there is another important class of solid-state sensor (and camera): this is the linear array sensor or line-scan sensor. In effect, these sensors are simply a one-dimensional array (row) of photosites, and use exactly the same technology as the two-dimensional array sensors. They differ in two important characteristics, however. Firstly, these sensors can have between 256 and 4096 photosites in a row and, hence, can achieve much greater resolution than state-of-the-art array cameras. Since they are inherently one-dimensional devices they can only take pictures of slices of a two-dimensional scene, and if a two-dimensional image is required, several such slices must be acquired. Thus, these sensors are best suited to the inspection applications in which the scene to be scanned is in continuous linear motion (or, indeed, where the camera itself can be translated). The second point to notice about these sensors is that the video signal that they generate does not correspond to any particular video standard and what is produced is essentially a time-varying analogue voltage which represents the incident light along the line of photosites. The repercussion of this characteristic is that most systems which use a line-scan camera tend to have custom-designed computer interfaces, although it is worth noting that matched line-scan cameras and digitizers and, indeed, line-scan digitizers are now commercially available (e.g. from Datacube Inc. and Imaging Technology Inc.).

2.2.3 Camera interfaces and video standards

The monochrome video signal standard used in the United States and Japan is RS-170, a subset of the NTSC (National Television Systems Committee) standard. Europe uses the international CCIR (International Radio Consultative Committee) standard, which is similar to, but incompatible with, RS-170. Television video signal standards are, unfortunately, inappropriate for the requirements of machine vision and both standards present essentially the same problem to machine vision applications. However, since the entertainment industry is still a far more lucrative market for camera manufacturers than machine vision is, few image sensors and cameras deviate from television standards. We will have a brief look at the CCIR standards to better understand the limitations it imposes on the imaging system.

Approximately fifty pictures per second are necessary for a flicker-free television picture. As the generation of a complete new picture every $\frac{1}{50}$th of a second (20 ms) would place a severe load on the capabilities of the video amplifier so, in an attempt to obtain a flicker-free image, the odd lines of a picture are scanned first and then, in the following $\frac{1}{50}$th of a second, the even lines are scanned. This means that a complete picture is established in $\frac{1}{25}$th of a second. This procedure whereby every alternate line is skipped in each field is called *interlaced scanning*. Thus, the number of complete pictures per second, i.e. the *picture frame frequency* is 25 Hz; the *raster field frequency* is 50 Hz. Each frame comprises 625 lines in the CCIR standard (525 in the NTSC RS-170 standard) thus, with twenty-five pictures of 625 lines, $625 \times 25 = 15\,625$ lines will be scanned in a second. The so-called *line frequency*, f, is, hence, 15 625 Hz.

The signal from a line comprises the picture information, a synchronizing pulse marking the end of a line (horizontal sync. pulse), and a blanking period. The time available for one line is:

$$T = 1/f = 1/15\,625 \text{ s} = 64 \times 10^{-6} \text{ s} = 64\ \mu\text{s}$$

11.5 μs of this are used for the blanking and synchronizing signal. This consists of a 6.5 μs *porch* blanking signal and a 5 μs sync. The sync. pulse signals the beginning of one line and the end of another and the porch represents a quiescent voltage level which prevents the beam appearing as a bright line during the line flyback (when being displayed on a video monitor).

The end of the picture consisting of 625 lines, i.e. the frame, is characterized by several picture pulses which are significantly different from the line pulses. During these pulses, the picture is blanked and the beam returns to the top of the picture for the next frame. A (simplified) diagrammatic representation of a video signal is depicted in Figure 2.4.

Unfortunately for machine vision, the CCIR standard specifies a 4:3 horizontal-to-vertical aspect ratio for video signals. The television monitor has a similar aspect ratio and thus an image of a square will appear square. However, as we shall see shortly, image acquisition devices often have a 1:1 aspect ratio, meaning that the rectangular video picture is effectively squeezed into a square with a consequent geometric distortion. This has very severe repercussions for vision-based gauging applications which we will address in Chapter 3.

2.2.4 Characteristics of camera sensors

There are several measures of the usefulness and reliability of a camera sensor; these characteristics will be briefly considered in turn.

☐ *Resolution*

Given that the purpose of the camera sensor is to convert an incident analogue optical image to some electrical signal, the resolution of a sensor can be defined as the number of optical image elements which can be discriminated by that sensor

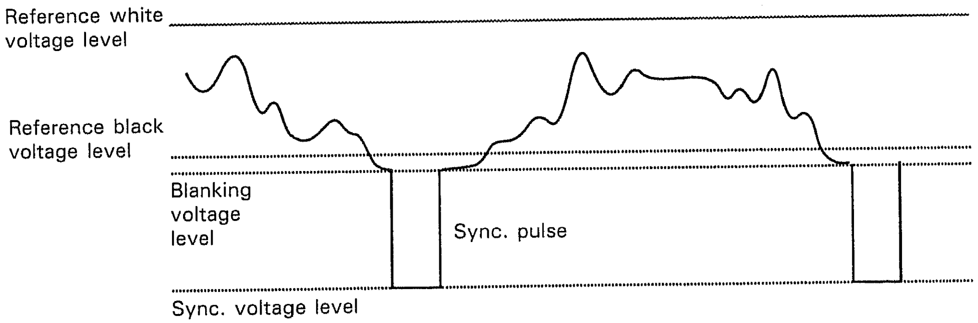

Figure 2.4 Composite video signal.

and represented by the resulting signal. The resolution is limited by the number of photosites in the sensor since this defines the frequency with which the optical image is sampled. The concept of image sampling is an important one and we will defer detailed discussion of this issue until Chapter 3. For the moment, note that in solid-state cameras the effective resolution of the sensor is approximately half that of the number of photosites in any given direction. Resolution is also limited by the array geometry and by how much opaque material separates the photosites: interline transfer sensors have less effective resolution than frame transfer sensors due to the presence of shielded buffers between each photosensitive line.

As we will see in Chapter 3, the resolution of the camera system is also constrained by the limitations imposed by the CCIR video standard and by the sampling frequency of the analogue-to-digital converter which converts the video signal to a digital image.

Tube camera resolution is a function of the electron-beam diameter relative to the area of the photoconductive layer. Tube camera resolution is generally higher than that of solid-state cameras and easily outstrips the limitations imposed by the CCIR standard.

☐ *Geometric faults*

For television cameras with electron-beam scanning, deviations in constancy of the vertical and horizontal deflection show up as faults in the geometrical placement of the picture content. Standard industrial cameras are not designed as measuring cameras but to generate a picture for subjective human examination, and they exhibit relatively large geometrical faults. For the standard industrial television camera it is usually ±1 per cent or ±2 per cent of the picture frame and it is usually much larger with cheaper cameras.

With CCD cameras there are no geometrical faults due to electron beam scanning; any geometric distortion is due to the lens.

☐ *Sensitivity and transfer linearity*

The input signal of an image sensor is a distribution of light or brightness. The output signal is a current or voltage which is related to the brightness. The sensitivity of the sensor is defined as the ratio of the output magnitude to the input magnitude.

In general, the output will be some power function of the input magnitude:

$$\text{output magnitude} = (\text{input magnitude})^{\gamma}$$

where γ (gamma) is the exponent of the transfer function which, rearranging the above equation, is given by:

$$\gamma = \frac{\log(\text{output magnitude})}{\log(\text{input magnitude})}$$

A linear sensor would have a γ of 1. The following are typical values of transfer

linearity for common types of sensor:

Image sensor	*Gamma*
Sb_2S_3 vidicon	0.6
PbO plumbicon	0.95
CCD	1

☐ *Lag*

Lag is often defined as the percentage of the signal current at a certain point of the target some period after the illumination has been switched off. Typical values for an elapsed time of 60 ms are shown below:

Image sensor	*Lag*
Sb_2S_3	20%
PbO plumbicon	2%–5%
CCD	1%

The lag introduces a limitation in the practical industrial application of TV cameras in terms of the permissible speed of movement of an object under consideration.

☐ *Spectral sensitivity*

The spectral sensitivity of an image sensor system is defined as the variation of the output as a function of the wavelength of the incident light. Referring to Figure 2.5, we see that solid-state image sensors are sensitive to light in the near infra-red region of the spectrum, and since such radiation does not carry any useful visual information, an infra-red cut filter should be used with such cameras, particularly if they are to be used with incandescent lighting.

☐ *Blooming*

If the sensor of a television camera is subjected to intensive brightness, then the

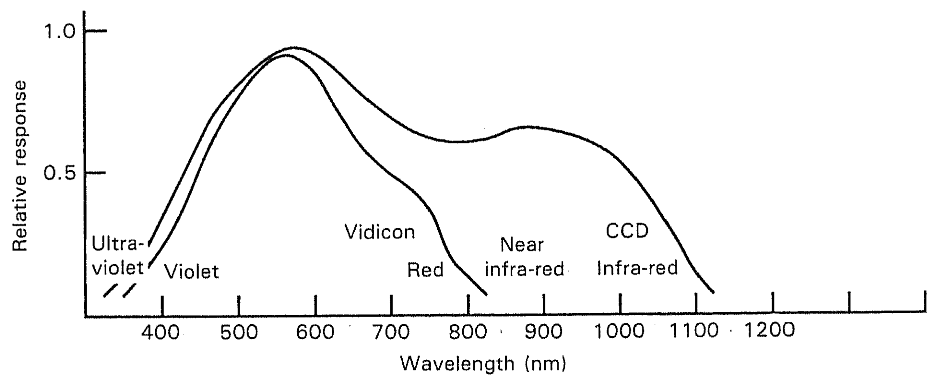

Figure 2.5 Spectral sensitivity of Vidicon and CCD cameras.

excess charge carriers spread into neighbouring zones and light is registered there also. Thus, a thin beam of very bright light will be sensed over an area considerably larger than the cross-sectional area of the beam. This effect is called 'blooming' and is especially noticeable with specular (mirror-like) reflections. While tube cameras are most susceptible to blooming, solid-state sensors, too, will exhibit this characteristic under extreme conditions.

☐ *Noise*

Noise, unwanted interference on the video signal, is defined quantitatively by the signal-to-noise ratio (S/N), i.e. the ratio of the amplitude of the wanted signal to the average amplitude of the interference. For television cameras, the signal-to-noise ratio is defined as the peak-to-peak video signal divided by the effective amplitude of the noise. If one follows general practice in telecommunications and expresses the ratio of electrical variables of the same unit in logarithmic terms, one can compute the level, defined as twenty times the decimal log of the ratio of the linear variables and expressed in decibels (dB). Thus, a signal-to-noise ratio of 20 dB means that the picture quality is very bad (20 dB implies that $\log_{10}(S/N)$ equals 1 and, thus, the signal-to-noise ratio is 10:1). For satisfactory quality, a signal-to-noise ratio of 40 dB or more is required.

Table 2.1 CCD Camera systems

Vendor	*Model*	*Camera type*	*Sensor resolution*	*Output signal*	*Interface required*
Sony	XC-77CE	CCD Area	756 × 581	CCIR	Standard framestore
Sony	XC-57CE	CCD Area	500 × 582	CCIR	Standard framestore
Panasonic	WV50	CCD Area	500 × 582	CCIR	Standard framestore
Pulnix	TM-540	CCD Area	500 × 582	CCIR	Standard framestore
Fairchild	CCD3000	CCD Area	488 × 380	CCIR	Standard framestore
Hitachi	KP-120	CCD Area	320 × 240	CCIR	Standard framestore
Videk	Megaplus	CCD Area	1320 × 1035	Non-interlaced analogue; digital	Special framestore
Fairchild	CCD1600R	CCD Line-scan	3456 × 1	Analogue	Requires camera controller
Honeywell	HVS 256	CCD Line-scan	256 × 1	RS232; RS422	Conventional serial port
Fuji	PJ1	CCD Line-scan	2048 × 1	Analogue; Digital flags	Can be directly interfaced to PLC[a]
Analytic Vision Systems	IMS-90	CCD Line-scan	1024 × 1 2048 × 4 4096	Analogue; Digital	Slow-scan framestore; can be directly interfaced to PLC

[a] PLC: Programmable Logic Controller.

2.2.5 *Commercially available cameras*

Table 2.1 summarizes the characteristics of a number of popular CCD cameras which are available at time of writing. We have listed only CCD cameras since these are the most popular for industrial machine vision.

Exercises

1. Identify the basic operational differences between vidicon-tube cameras and solid-state cameras. Which type of camera is more appropriate for use in FMS (flexible manufacturing systems) requiring automated vision for either inspection or robot guidance? Why is this so?
2. What limitations does the European CCIR television standard impose on the effective resolving power of imaging equipment? Is it an appropriate standard for machine vision? Explain.
3. It is required to image a 5 cm wide object using a CCD camera, whose sensor measures 8.8 mm × 6.6 mm, at a distance of 0.5 m. What focal length lens should you choose? What options do you have if you wish to image an object which measures only 1 mm × 1 mm.

References and further reading

Batchelor, B.G., Hill, D.A. and Hodgson, D.C. (eds) 1985 *Automated Visual Inspection*, IFS (Publications) Ltd, U.K.

Blouke, M.M., Corrie, B., Heidtmann, D.L., Yang, F.H., Winzenread, M., Lust, M.L. and Marsh, H.H. 1987 'Large format, high resolution image sensors', *Optical Engineering*, Vol. 26, No. 9, pp. 837–43.

CCIR, International Radio Consultative Committee 1982 *Recommendations and Reports of the CCIR, 1982*, XVth Plenary Assembly, Geneva, Vol. XI, Part 1.

Dunbar, P. 1986 'Machine vision', *Byte*, Vol. 11, No. 1, pp. 161–73.

Herrick, C.N. 1972 *Television Theory and Servicing*, Reston Publishing Company, Virginia.

Market Intelligence Research Co. 1986 *Solid-state Cameras for Machine Vision*, Market Intelligence Research Company, 4000 Middle Field Road, Palo Alto, California, 94303, USA. (*Vision*, SME, Vol. 4, No. 1, p. 12.)

Tassone, J. 1987 'An illumination system for inspecting printed circuit boards with surface mounted components', *Vision*, SME, Vol. 4, No. 4, pp. 1–5.