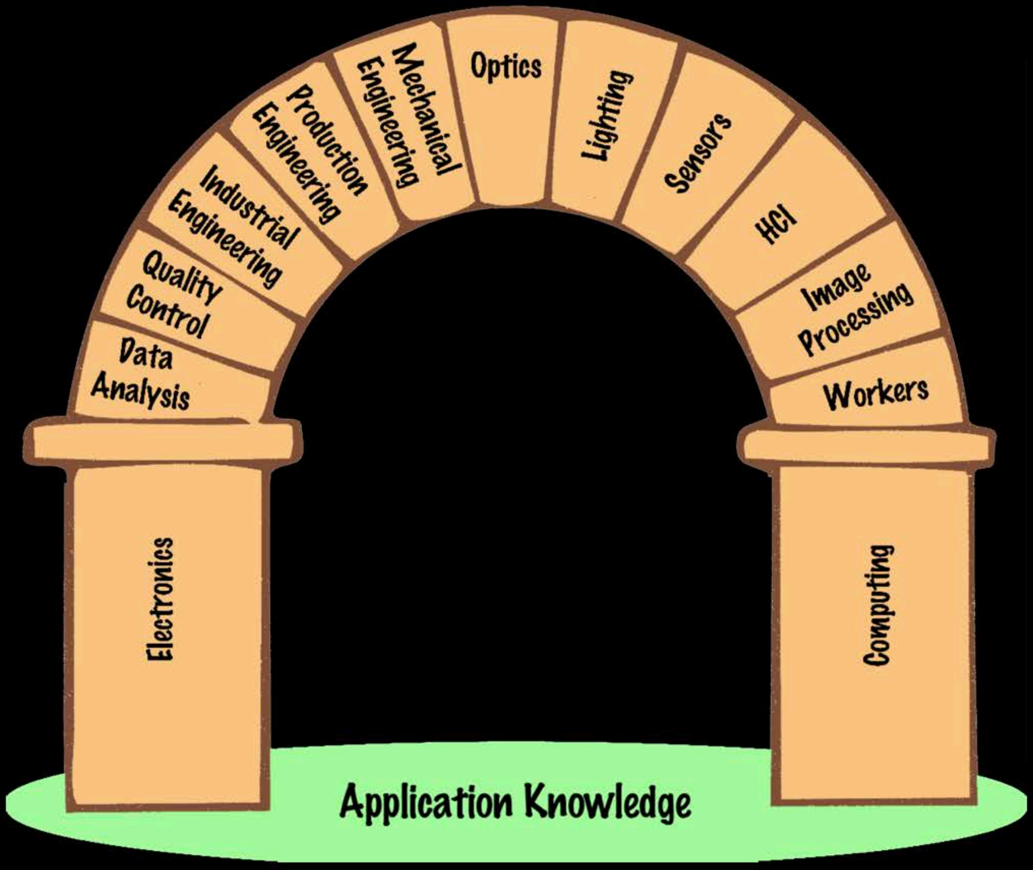
Chapter 2

A New Vision



Terms Used

Machine Vision (**MV**) The use of electronic/computer devices for optical, non-contact sensing to receive and interpret an image of a real scene automatically, in order to obtain information and/or control machines or manufacturing processes.

Machine Vision System: A complex assembly that is able to sense its environment optically. It incorporates mechanics, electronics, optics and software, which must all be integrated to function together harmoniously.

Image Processing: Mathematical operations on a digital image to transform it into another, image, to translate it into some more convenient format, or to take measurements on it.

Digital Image: Digital representation of an image. In this book, it is taken to be an array of numbers. Other mathematical represtations are possible.

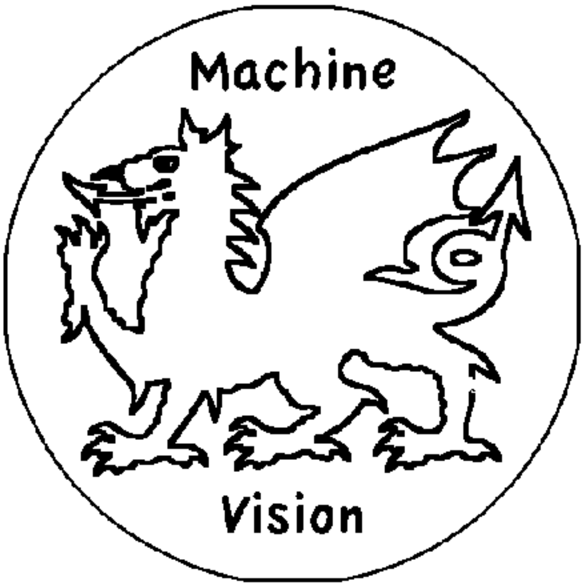
Pixel
An element in a digital image, representing the brightness/colour, of a single point in a real scene.

Motivation

We saw in the previous chapter that there are many very different models for natural vision and that human vision has many subtle features that we do not need to emulate. We turn our attention now to Machine Vision (MV), which is often very different from anything found in nature. Much of what follows is about Machine Vision Systems (MVSs) that are designed to inspect objects during, or shortly after, manufacture, or to control machines in a factory. For these purposes, an MVS does not need to be able to see in the same way that a person or an animal does. The aim of Chapter 1 was to free us from the notion that human beings have the only model of vision that works effectively. Here we will go further: we will not be confined by the idea that Machine Vision should necessarily match human or animal vision.

Before we begin in earnest, it is necessary to explain why so many of the illustrations in this book are based on industrial artifacts. There are several reasons:
Health, safety, work, and enjoyment of modern life all rely on the correct construction and operation of manufactured items. However, their failure can be frustrating, inconvenient, even dangerous, costly, and wasteful of both money and materials. Faulty components and mistakes during manufacture can result in significant dangers in many products. Of particular concern are food, automobiles, aircraft, electrical devices, pharmaceuticals and health-care devices.

-



Blank page

Here are some more reasons for taking an interest in industrial applications:

- Human beings cannot always detect faults reliably at the speeds needed to match modern manufacturing methods. There is a huge number of potential industrial applications for machines that can inspect and measure industrial artifacts automatically. These range in complexity from trivial to extremely difficult. A car provides a good example. A single vehicle has about 20,000 components, hundreds of which are safety critical and require careful inspection.
- From an educationalist's and technical author's point of view, illustrating Machine Vision with reference to industrial products is very attractive, because the subject can be addressed at many different levels of complexity. It is possible to select the application to illustrate a general idea in the most appropriate manner.
- Many people have repetitious, tedious and boring jobs. Some tasks expose workers to danger and long-term health risks. In many situations, Machine Vision can eliminate those dangers (How this can be reconciled with the need to provide work for people of all abilities, is beyond the scope of our present discussion.)
- Many factories with high levels of automation have relatively few shop-floor personnel.(**Figure 2.1**)

Figure 2.1 The major features of an automated food manufacturing line are high production speed, low unit value, the necessity for cleanliness and product safety. There are very few personnel in the factory. Repetitive inspection tasks, such as this, cannot be performed reliably by human beings for more than a few minutes.

Providing effective automated inspection procedures is both essential and urgent. In many situations, the great speed and precision of modern manufacturing processes demands far greater capability than any human capability. Human inspectors just cannot cope with situations like that shown in **Figure 2.2.**

You may well have seen evidence of the existence of an industrial MVS on television. When watching a business-news reporter standing in a factory, you might see a small area in the background production-area that is very brightly lit. (The light may be flashing.) That is almost certainly a sign that a vision system is at work.

Figure 2.2 If Charlie Chaplin were to remake the film "Modern Times", he might well incorporate scenes like this. The unending and overwhelming stream of bottles has a night-mare quality, even though we cannot hear the noise of the machines.

Machine Vision System

While we will often refer to Machine Vision, we should properly use the term *Machine Vision System* (*MVS*) because we cannot isolate the *process* of "seeing" from the *equipment* needed to do that. A Machine Vision System combines several different technologies as the picture of the arch at the front of this chapter indicates.

Figure 2.3 shows a human inspector at work. This provides a useful way to understand what a Machine Vision system is and how it operates. Substitute a camera for the inspector's eyes and a computer for his brain. His arm is replaced by a simple *accept/reject* mechanism for rejecting faulty widgets from the conveyor belt.

Figure 2.3 Elements of a typical Machine Vision system translated into human terms. Notice the following: object transport mechanism (conveyor belt), dedicated lighting, optics (spectacles), image sensor (eyes), product information database (book), data processing unit (brain), device for rejecting faulty product (left arm), reject bin.
Original caption: "Checking the spelling of seaside rock". *Rock is a confectionary bar, typically with the name of a holiday resort spelled out by lettering that runs internally along its whole length.*

A Machine Vision System is far more than a camera interfaced to a computer. **Figure 2.4** shows the very limited level of understanding that can be achieved when explaining Machine Vision, for example, at a dinner party, but it is far from complete.

Figure 2.4 This naive set-up is adequate only as a "dinner party" description of a Machine Vision system; it bears little resemblance to a practical installation designed to survive in a dirty factory. Principal deficiencies:

- *Unspecified lighting*
- *No protection for the camera and lens.*
- *No indication how objects are presented for viewing.*
- *No way to synchronise image capture*
- *Unprotected computer*
- *Unproven reliability of software*
- *Algorithm robustness has not been established*
- *Unspecified human-to-computer interface (HCI)*
- *Unspecified action when a defect is found*

Figure 2.5[L] shows a popular misconception: a camera simply aimed at the product to be inspected. There is no protection for the camera, which is exposed to many dangers: accidental and malicious. By this time, you might reasonably expect to find a picture of a real Machine Vision system and we will not disappoint. Be aware, however, that Machine Vision systems are distinctly non-photogenic. Figure 2.5[R] give a better view of a vision system but shows a mysterious stainless steel box. Even this is unsatisfactory because it gives no hint of the high level of engineering needed to ensure that the system is robust and reliable. Surely this one of the most boring pictures you have ever seen? The interest lies in what it can do.

Figure 2.5 Vision systems are not photogenic. (Left) The naive approach: simply place a camera and a light source (LED ring light) beside a production line. (Right) More robust system for inspecting flat food products. The stainless steel enclosure is designed to protect the equipment from food contamination and vice versa. It houses the camera, lights, various optical devices and a computer. The enclosure also reduces the effects of variable ambient lighting. The conveyor belt is an integral part of the system. To maintain safety and good hygiene, the machine must be water- and dust-proof but be easy to strip down and clean.

Before we move on, there is an important point to note: factories are dirty places. (**Figure 2.6**) in addition, there are many fingers itching to fiddle with cameras and lenses, even steal them. Protecting the camera lens and light source in an MVS is therefore of great importance. That is why a rugged box is needed to protect them. (**Figure 2.5[R]**)

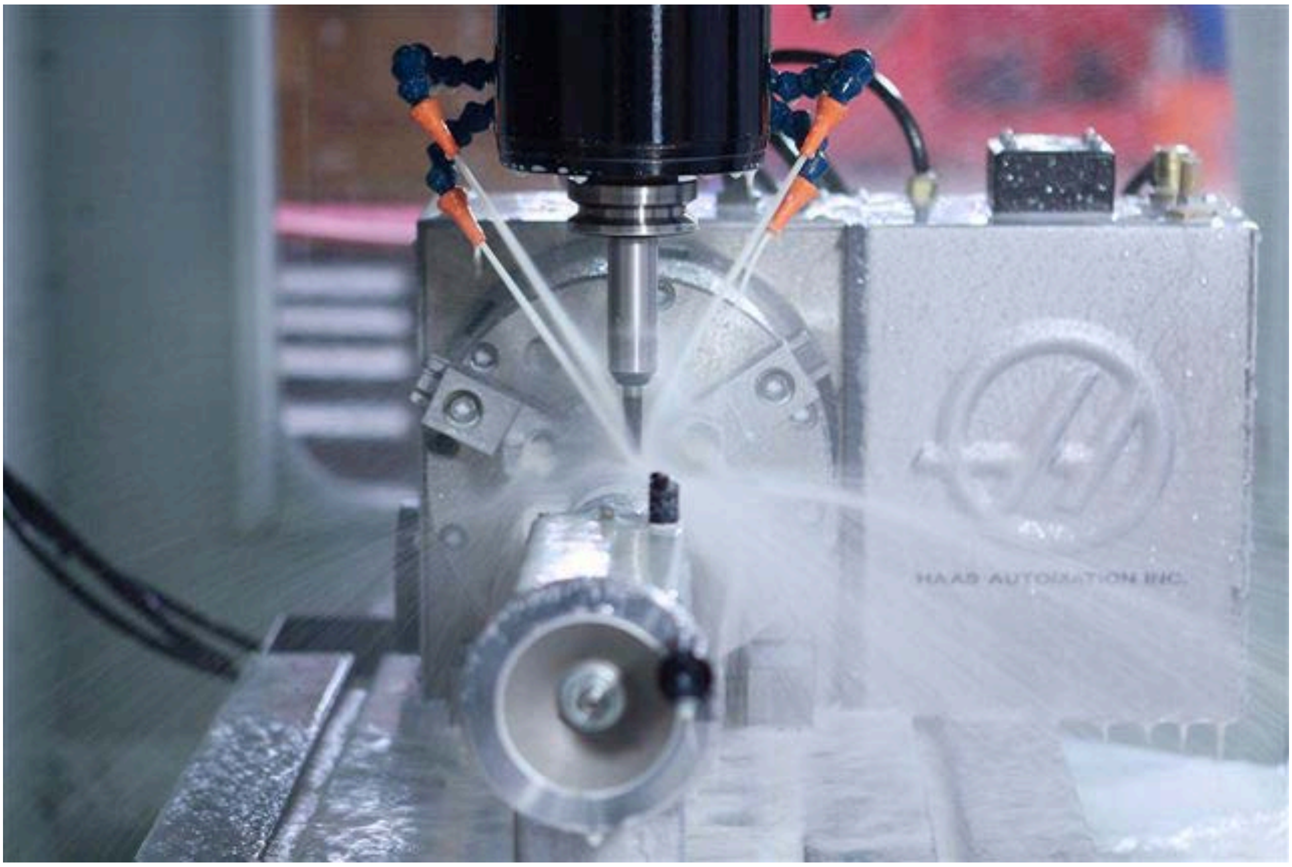


Figure 2.6 Factories are sometimes very dirty places! Placing an unprotected camera and lights close to a machine tool such as this lathe would be foolish. Apart from the obvious cloud of cutting fluid, numerous tiny droplets are created in the atmosphere and can quickly contaminate optical surfaces, degrading image quality.

The protective enclosures used for traffic enforcement cameras ("speed cameras") provide a more familiar example. (**Figure 2.7**) Industrial MV cameras are often built like this, to withstand tampering, as well as violent accident and attack.

Figure 2.7 Lessons for Machine Vision. (Left) Radar measures vehicle speed. (Right) Road markings enable speed measurement. High contrast, caused by shadows, makes viewing difficult. (In an industrial installation, the viewing environment is under better control.} Other major points of note are as follows:

- *The camera is protected inside a rugged, tamper-proof, weather-proof enclosure.*
- *The lighting is controlled, using a bright flash lamp.*
- *The scene being viewed has been adjusted to make analysis easier, by adding road markings*
- *Image capture is triggered by an external sensor: the radar speed measurement.*
- *The camera and light source are carefully placed relative to the object of interest (i.e. the rear end of a receding vehicle, to maximise the intensity of the retro-reflected light from the registration plate)*
- *The field of view of the camera is carefully chosen.*
- *The camera is held rigidly in place, to minimise vibration, which would produce blurred images.*

What Is in the Box?

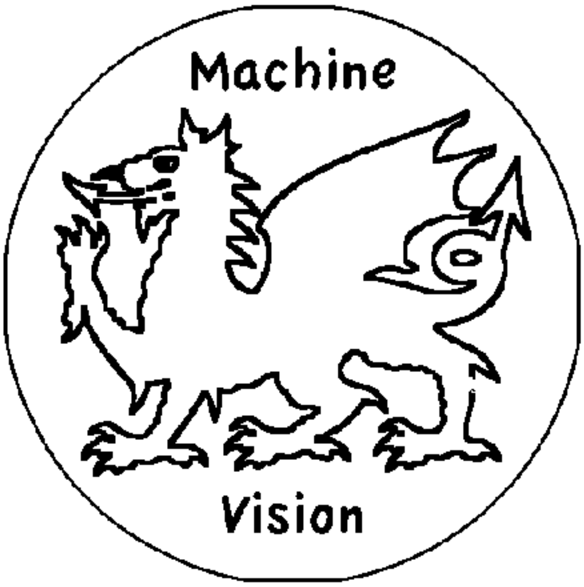
Machine Vision systems for industrial applications invariably require the design and construction of devices, integrating:

- Mechanical handling
- Lighting
- Optics
- Camera(s)
- Dedicated electronic hardware (Sometimes, this is not needed.)
- Computer hardware
- Software (This is based, in part, on mathematical procedures for processing and analysing images. These are called algorithms*.)*
- Human-to-computer interface (HCI in the picture at the front of this chapter)
- Machine-to-machine interface

Furthermore, these must all be made to work together harmoniously.

Machine Vision does not normally attempt to emulate either human or animal vision. The requirement that an MVS must be fast, cheap and reliable will normally be awarded precedence over any attempt to copy nature as doing so would usually be an unrewarding distraction.

Designers of Machine Vision systems (Vision Engineers) enjoy one very great advantage: production staff in a factory know, at every given moment, what kind of widget is being made. So, it is possible to anticipate, in broad general terms, what kind of image a vision system should expect to see.



Blank page

Any major deviation from the image associated with"good" product is indicative of a faulty product. We may restate this as the Law of Anticipated Appearance:

> *If the present widget does not look like "good" product, then it is faulty (and should be scrapped or reworked).*

Or, in more colloquial terms

> *"If it ain't what it is supposed to be, it is wrong."*

Exploiting this and other so-called application knowledge requires that a vision engineer has an awareness and understanding of a wide range of subjects, especially optics computers and electronics. The multi-disciplinary nature of Machine Vision is emphasised in the arch pictured at the front of this chapter and **Figure 2.8**. Let us now consider each of these supporting technologies in turn.

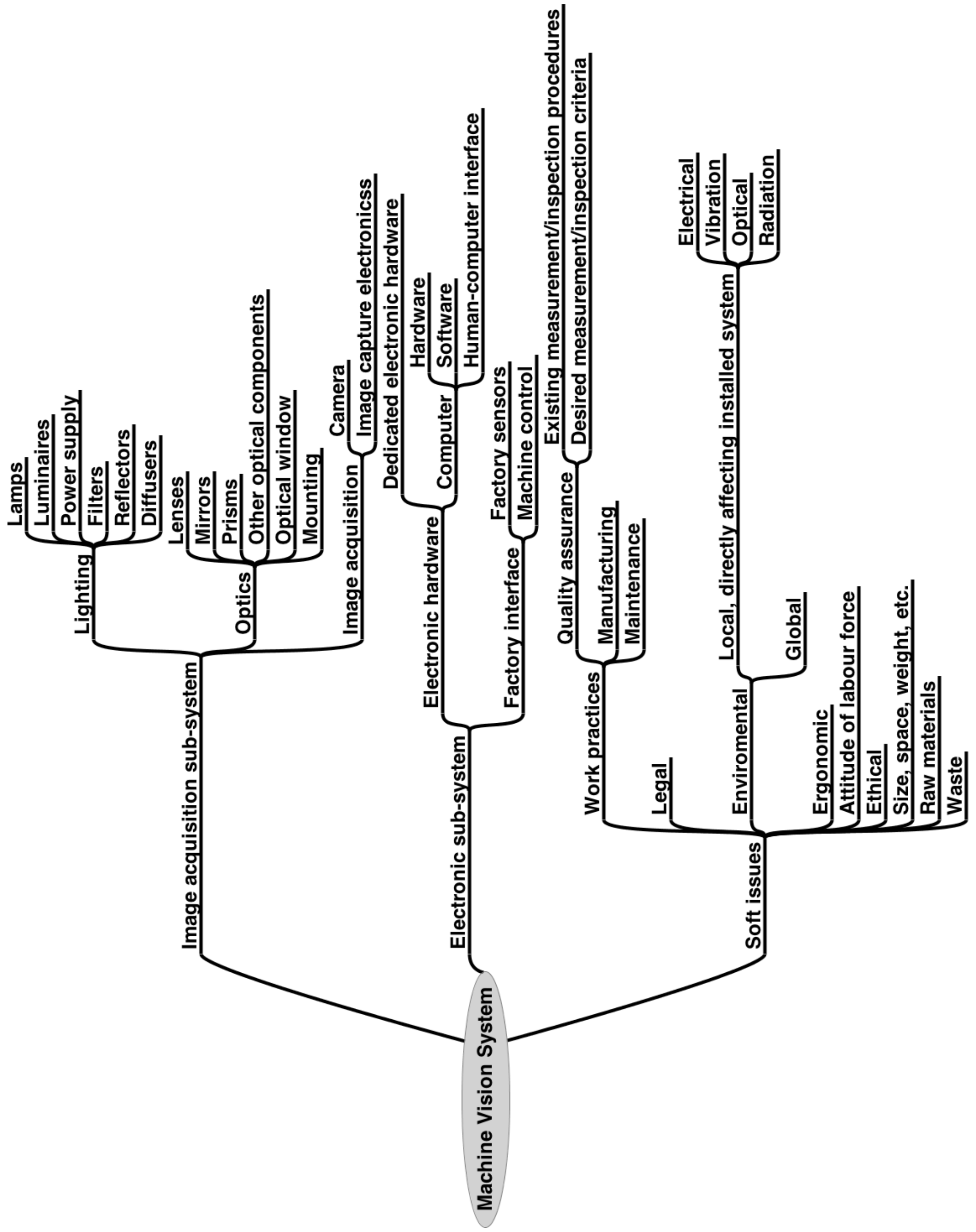


Figure 2.8 Viewing Machine Vision as a sub-set of Systems Engineering. For the sake of clarity, this is an abbreviated version of the real-life concerns of a Vision engineer.

Mechanical Handling

Human inspectors can often accommodate a variable viewing angle and will turn a widget around to obtain a good view of each critical feature. However, for a vision system to be most effective and reliable, widgets should always be held and observed in the same way, from fixed points. (**Figure 2.9**) Mechanical handling is necessary to move a widget towards a camera, hold it there, while the camera captures an image and then move it away again. Mechanical handling is also needed to move defective widgets off the production line and place them in a reject, or rework, receptacle.

What mechanical handling device is needed depends upon the object/material to be examined. It may, for example, be a simple flat-top table, a specialised clamp, a conveyor belt, or a multi-axis robot. Many factories use a continuously moving conveyor belt to transport moving flat strip products and discrete objects. We will see in a little while that there is a special type of camera, known as a line-scan camera, that is ideal for viewing objects on a conveyor that is moving continuously. On the other hand, a normal camera, known as an array camera, will normally produce a blurred image in this situation. More of this later.

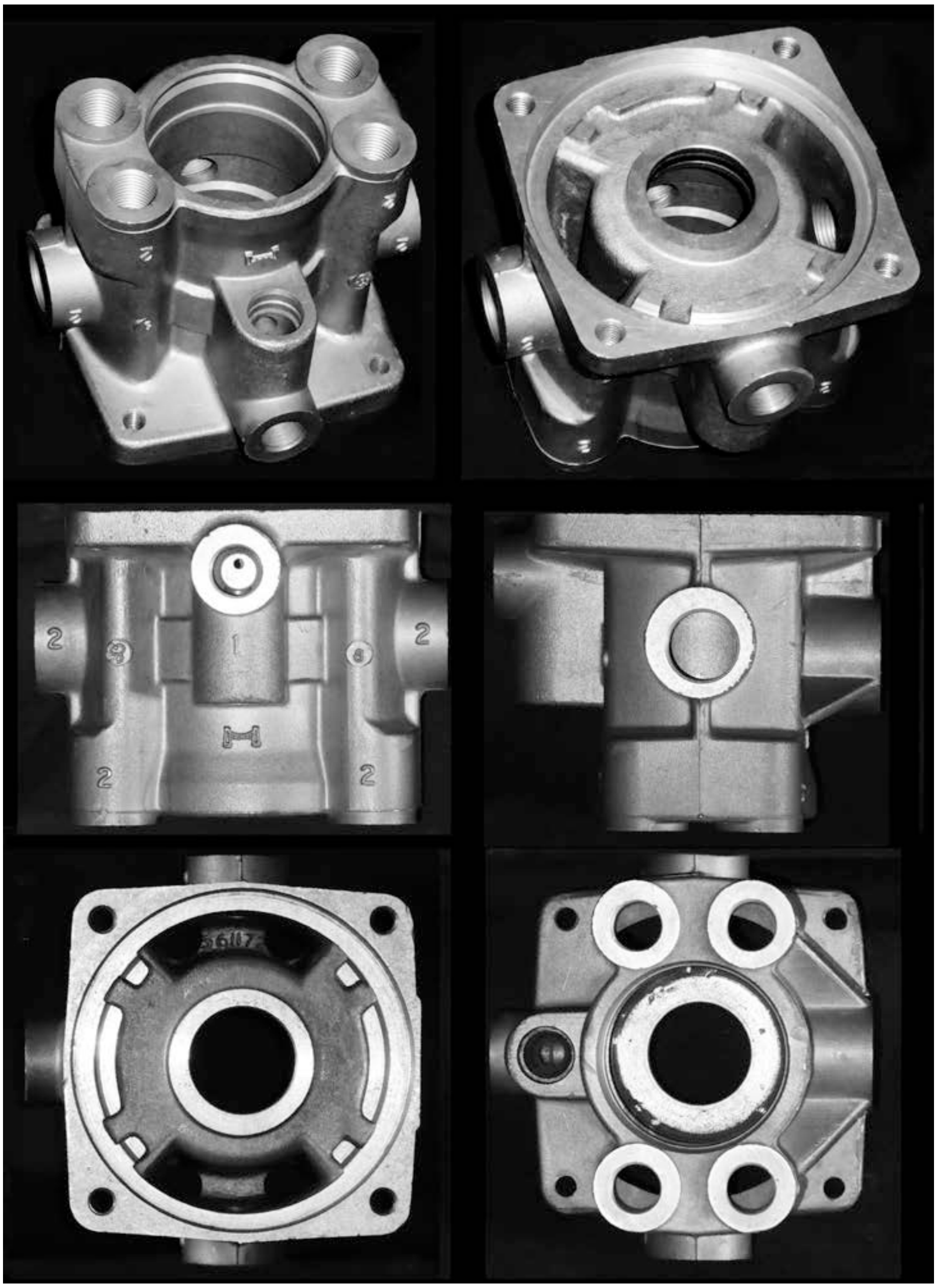


Figure 2.9 Views of the same object from different vantage points. [Alloy casting for an aircraft hydraulics system] (Top) Photographers usually prefer to show an object from an arbitrary view point that allows three faces to be seen obliquely. However, this is almost invariably ill-suited for Machine Vision. (Centre & Bottom) Four orthogonal views allowing more precise measurements

Figure 2.10[TL] shows an inspection station in which there are two cameras viewing bottles on an indexed rotating table: the table rotates (by 60° in this case), stops for a fraction of a second, the camera captures an image and the table rotates again. A conveyor belt, with a robotic parts feeder is used in the inspection system in Figure 2.10[BL]. Sometimes, it is necessary to use a robot to allow multi-angle viewing of an object. The camera and/or light source may be mounted on the robot, or on a fixture beside it. (Figures 2.10[TR] and [BR])

Figure 2.10 Automated visual inspection and robot vision scenarios.

Image Acquisition

Picture data has to be captured and stored within a computer before it can be processed. This is the responsibility of the so-called Image Acquisition sub-system, which itself has has six main elements responsible for:

Generating light

Optics

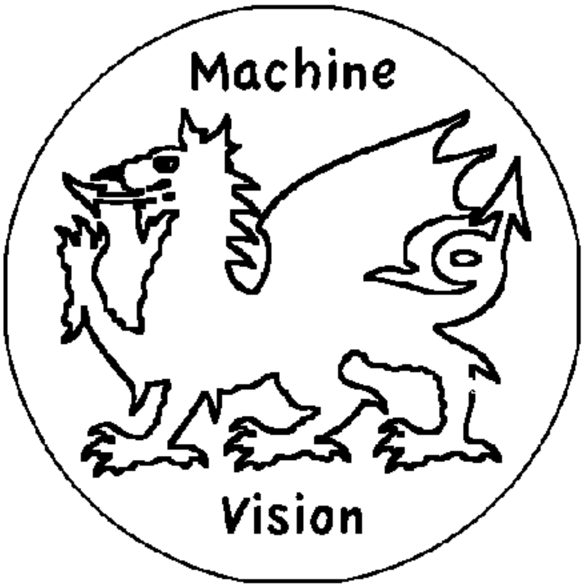
Guiding light onto the scene to be viewed

Collecting light from that scene, enhancing it optically and and forming an image on the camera's "retina"

Converting a pattern of light into electrical form (camera)

Conditioning the electrical signal to improve its quality

Converting the electrical signal into digital form.



Blank page

Generating Light

For many Machine Vision applications nowadays, LED (light emitting diode) illumination is much to be preferred over other methods of generating light. In recent years, there have major advances in the development of ultra-bright LEDs. These have several major positive features, most of which are not shared by other light sources:

An individual LED is a "point" source.

The light beam emitted by a single LED forms a narrow cone and can be directed where needed, simply by turning the LED, which is a small device, typically about the size of the head of a match.

Many LEDs can be assembled on a board, ring or other curved surface, or even on a flexible strip. **(Figure 2.11, 2.12, 2.13)**

LEDs use safe, low operating voltages.

LEDs are efficient in generating a lot of light in relation to the electrical energy used.

LEDs are available in a wide variety of colours. Each LED emits a fixed very narrow band of wavelengths (colours). LEDs emitting ultra-violet (UV) or infrared (IR) are also available. "White" LEDs can be made in one of two ways. In the first, light from a UV LED is directed onto a mixture of various fluorescent materials that together emit white light. The second method mixes light from red, green and blue LEDs to form what appears to the eye to be white light. (The spectrum of the mixed beam is not flat, as is normally expected for white light, but has three peaks. This can give rise to peculiar effects when illuminating some surfaces.)

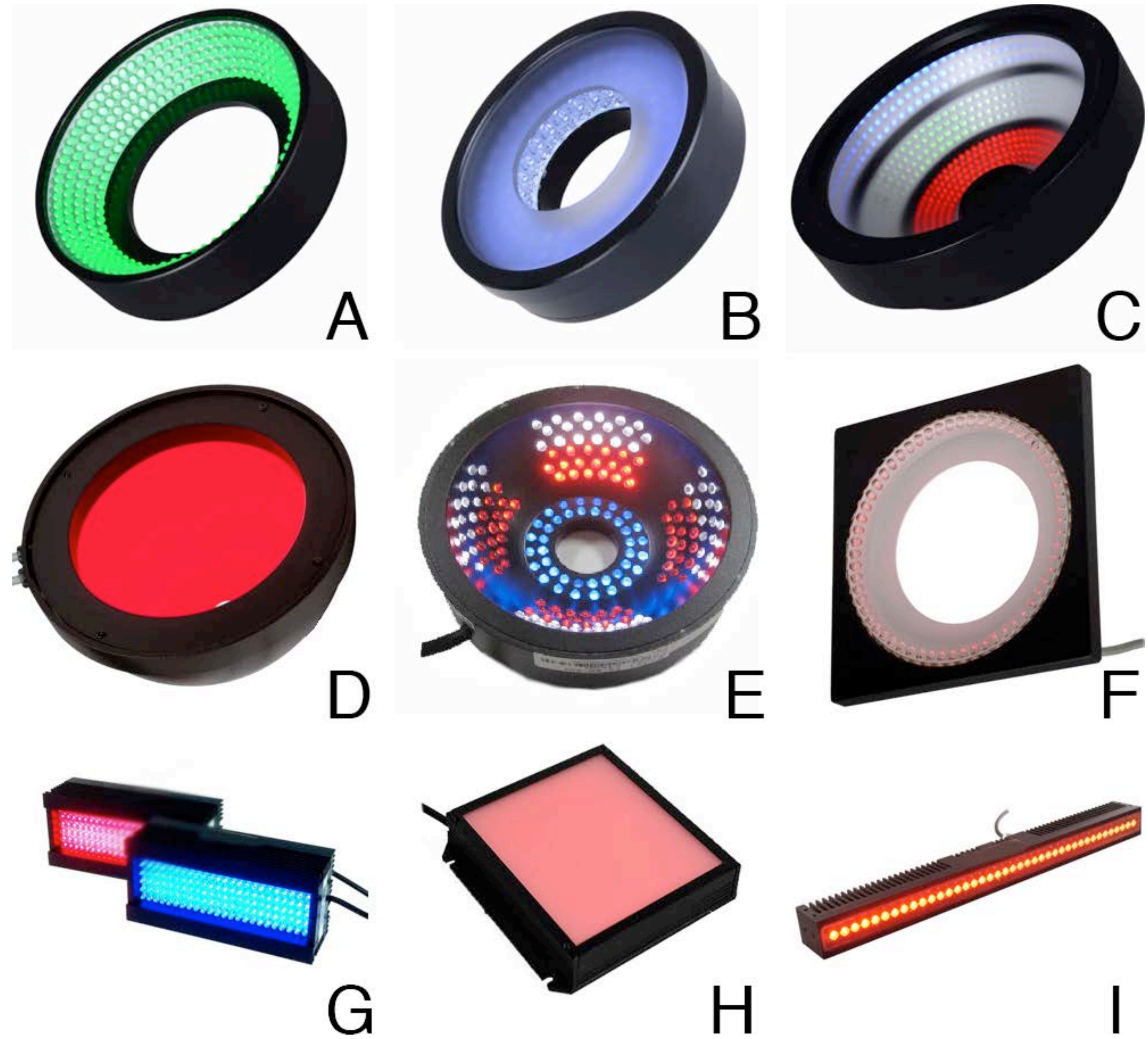


Figure 2.11 LED lighting devices designed for Machine Vision. (A) Projects light inwards at about 45°. (B) Fitted with diffuser. (C) Multi-colour. Each colour can be controlled separately. (D) Shadow-free dome illuminator. Camera can view through a hole at the centre. (E) As [A], Multi-coloured LEDs, can be controlled separately. (F) Projects light inwards, grazing ("dark field") illumination. (G) Small blocks of LEDS, various colours. (H) Uniform illumination, back lighting, other colours available. (I) Bar.

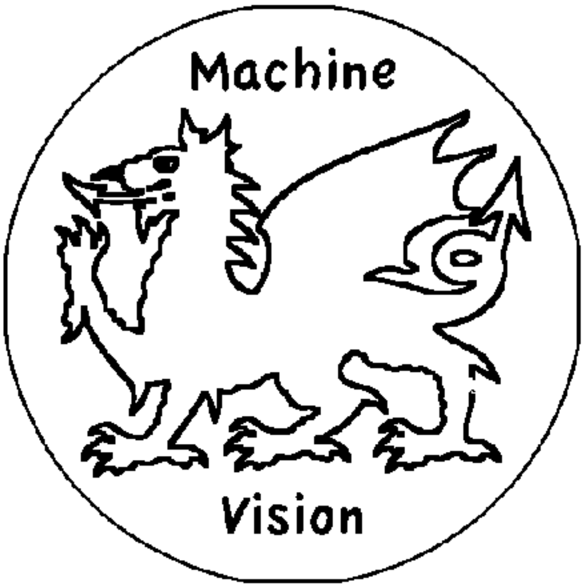
LEDs can be switched ON/OFF very rapidly (strobed). This enables them to be used in a stroboscope to "freeze" motion. It also enables their efficiency to be improved even further. (The use of strobed LEDs is not without danger: flashing lights can make rotating machines appear stationary and induce epileptic fits or migraine episodes.)

Many industrial applications of Machine Vision require close-up viewing in a well-controlled environment, in which case the benefits just listed are very helpful indeed. However, sometimes, it is necessary to use other light sources. For example, although they are much less efficient, old-fashioned, hot-wire lamps produce light that is close to sunlight in colour. Fluorescent tubes, of the type used for room lighting and more modern compact fluorescent lamps, require special precautions, otherwise horizontal stripes will appear across the digital image derived from a video camera.

There are several other options for lighting but the details need not concern us here. The point to note is that lighting is a specialised subject; we cannot for example, simply grab a torch or desk lamp, point it vaguely in the direction of the scene that we want to view and expect to obtain good images.Lighting and Viewing Arrangements

Major improvements in image quality can often be achieved by paying careful attention to the optical system. Apart from the initial cost of construction and installation, optical systems cost very little to run. Moreover, they are much faster than an electronic computer.

Figure 2.12 High-power LED lighting is used in theatres and for concerts. It is available in a range of colours. The luminaire on the left has three different coloured LEDs that can be controlled independently, by computer. Similar devices are available that provide red, green, blue, white, amber and ultra-violet light. LEDs can be strobed to overdrive them, thereby increasing the light output, as well as saving power and "freezing" motion. (Strobed lighting must be synchronised with image capture and care must be taken to guard people from flashing lights.) Both flood- and spot-light units are available. Some units allow automated control of the direction of the beam



Blank page

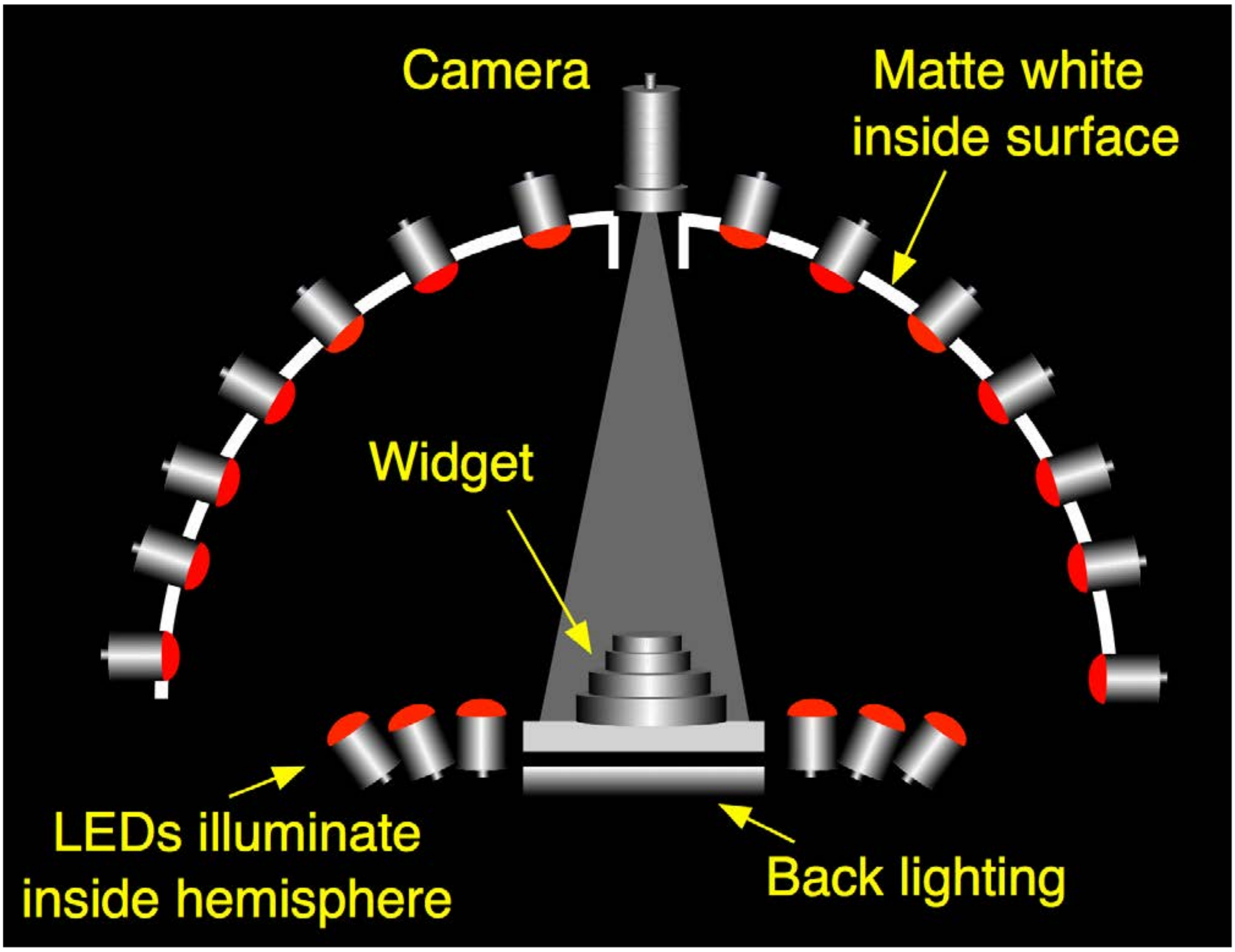


Figure 2.13 Flexible lighting-viewing arrangement for prototyping. For increased versatility, use multi-LED clusters, incorporating R, G, B & "white" devices. IR and UV LEDs can also be included. It can provide the following illumination: Grazing; Shallow-angle; Dark-field; Glint-free (for flat surfaces); Front; Coaxial illumination and viewing; Shadow- and glint-free (by lighting the inside of the hemispherical diffuser); Back illumination, Strobed illumination; Multi-view image capture, with switched lighting.

Colour Filtering

When trying to discriminate between colours, it is always a good idea to investigate whether optical filters can improve the contrast between them. **Figure 2.14** shows how effective even simple optical filtering can be. Sometimes, optical filtering is useful to make one colour disappear against a white background, or make colours merge so that they become indistinguishable. A colour camera incorporates three colour filters, which enhance the long wavelengths of light (red-orange), mid wavelengths (yellow-green-turquoise) and short mid wavelengths (blue-violet). These are known as RGB filters. However, very much better colour discrimination can be achieved sometimes, using a monochrome camera and one or more separate filters. Cheap colour filters use tinted glass, or gelatine, but the most startling results can be achieved using interference filters. These consist of multiple overlaid thin metal and dielectric (non-conducting) films and can achieve very sharp colour discrimination.

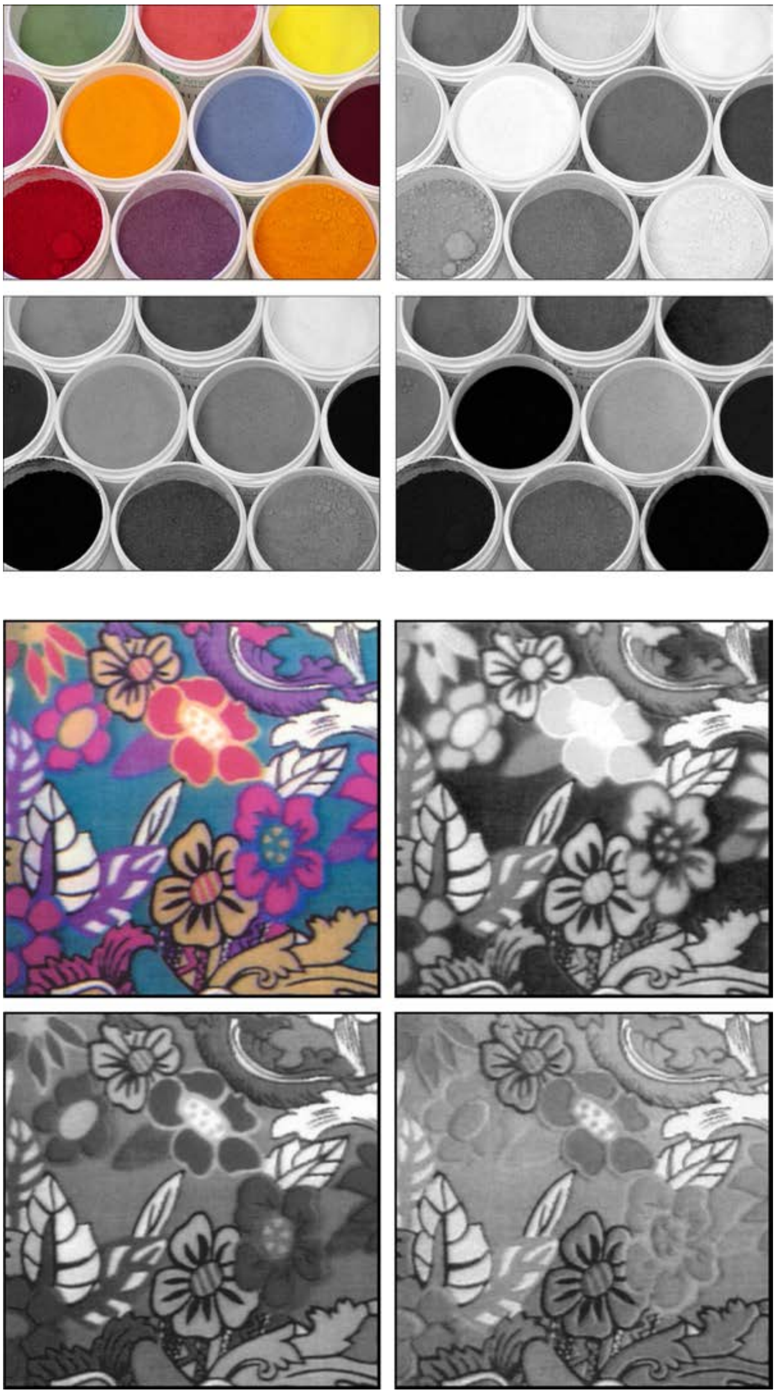


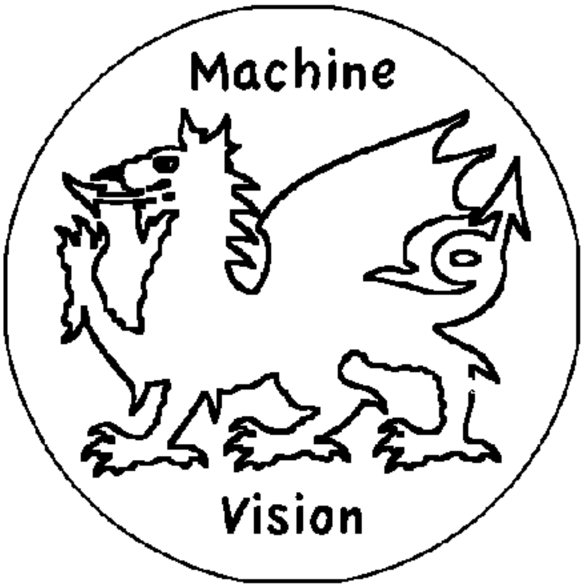
Figure 2.14 Two examples of optical filtering, using the R, G, B filters fitted within a standard colour camera. Notice some similar colours are separated by filtering, while some others become indistinguishable. *Layout:*

Colour image *Red (R-image)*
Green (G-image) *Blue (B-image)*

Colour samples: (Top) Powdered fabric dyes. (Bottom) Dress fabric

An optical filter can be placed immediately in front of the light source or camera; the effect is the same. (This remark does not apply if there is any fluorescence.) The filter must be chosen with the light source in mind. No filter can restore colour in an image derived from a scene illuminated by a sodium vapour lamp, typically used for street lighting. This type of lamp emit lights within a very narrow range of wavelengths (yellow). LEDs also emit light over a narrow part of the spectrum and this may mean that optical filtering is not needed. More will be said about this in Chapter 7.

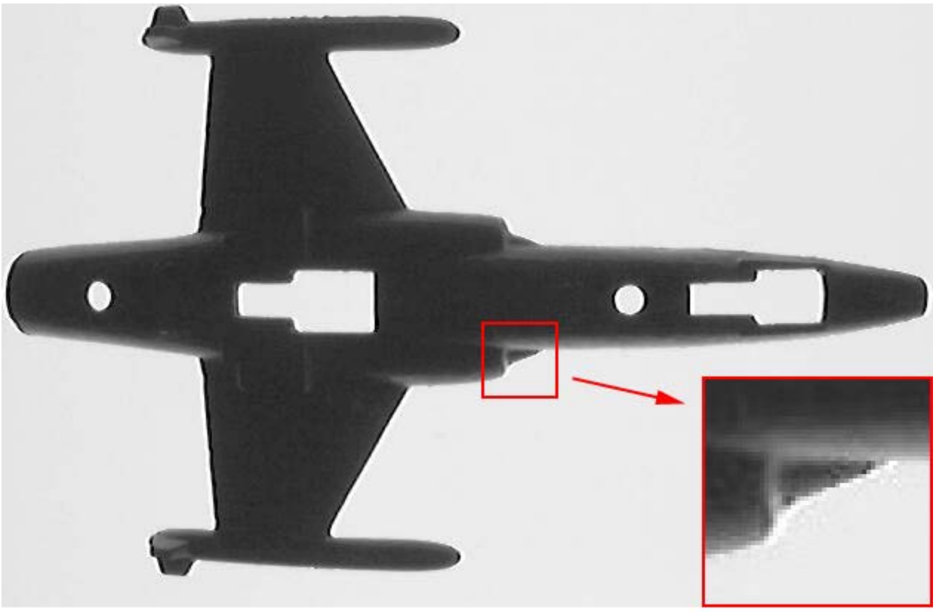
Photographers will realise that image quality can often be improved by blocking ultraviolet light (UV) by the addition of a filter. The same is true for infra-red (IR), which is particularly important as modern cameras are very sensitive in this part of the spectrum.



Blank page

Back-Lighting

Placing an opaque object between a broad-face light source and the camera is ideal for viewing silhouettes of opaque, non-reflective objects. This can also work well, even for transparent objects, such as glass or plastic bottles. **Figure 2.15** shows that back-lighting is not always successful when working with shiny objects. This difficulty can be avoided with care.



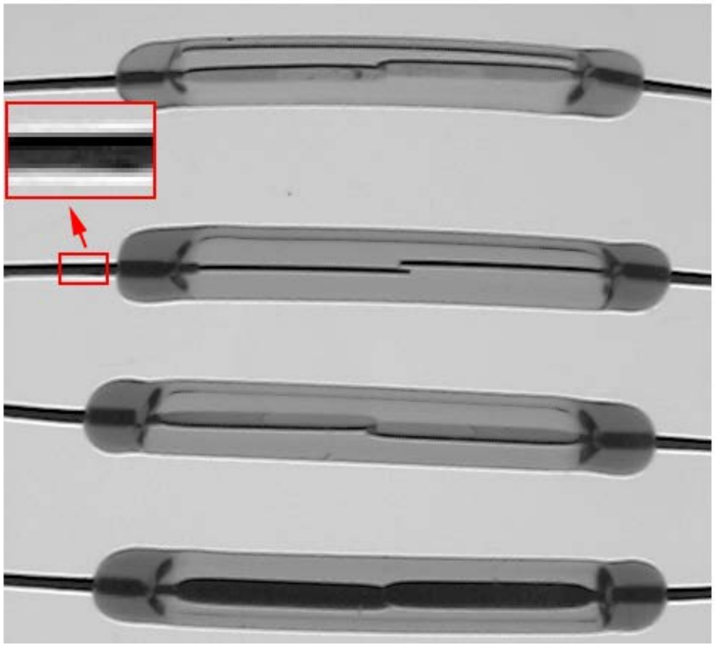


Figure 2.15 Back-lighting can produce very high contrast but there can be some difficulties. (Top) Zinc die casting. Notice the bright edges: the back-light source is too large. (Bottom) Reed relay inserts. The outer edge of the glass is well defined but there is a narrow bright "halo" above and below each wire.

Coaxial lighting and viewing

A coal miner wears a lamp on his/her helmet, so that wherever he/she looks is well lit, even though it may be very dark everywhere else. This does not create significant shadows. The trick is, of course, to place the light source very close to the miner's eyes. Similarly, the flash unit on a standard digital camera effectively illuminates the scene from a point very close to the lens. The result is a photograph in which even deep holes are well lit, again with only minor shadows By careful optical design, it is possible to do better than either the miner's head-lamp or the camera flash unit. The lighting-viewing arrangement in **Figure 2.16** shows the arrangement for coaxial illumination and viewing. Using this set-up, which incorporates a diffuser, a good quality image can be obtained from even bright shiny metal objects. The beam-splitter allows some light to pass straight through, while reflecting the rest. This arrangement completely avoids the camera from seeing any shadows, even at the bottom of a deep well. The slight offset of the miner's lamp and flash light from the view point does not achieve this perfectly, creating thin shadows.

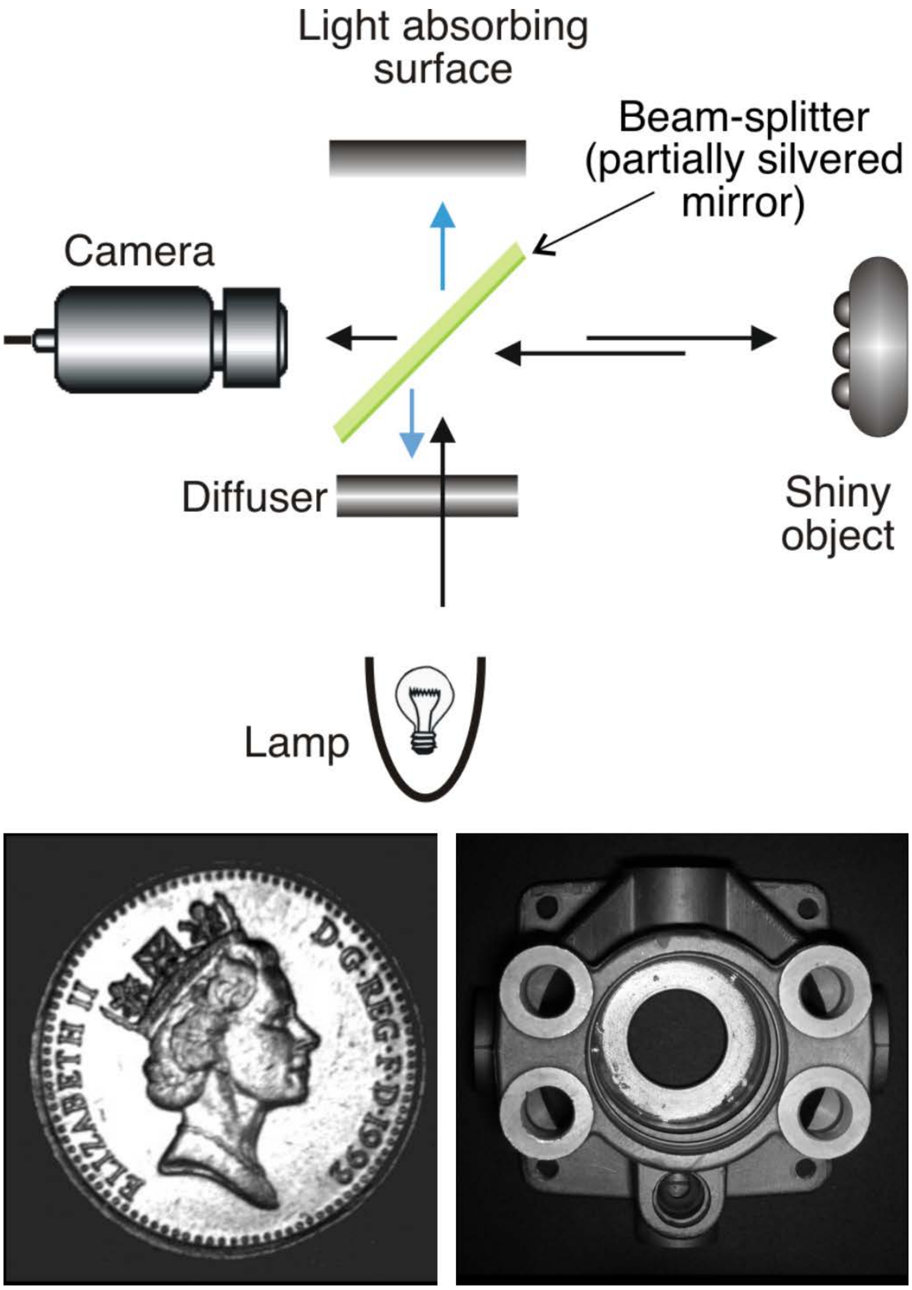


Figure 2.16 Coaxial illumination and viewing. (Top) Optical layout. Unwanted light is indicated by blue arrows. (Bottom-left) Sample image, bright "silver" coin. (Bottom-right) Cast metal component for an hydraulics system. Notice the absence of shadows even in deep crevices. The central feature is a well that is deeper than its diameter.

Front diffuse lighting

Glinting can be a problem when viewing bright shiny metal, objects. However, the difficulties can be alleviated by using a light source that has a broad face. **Figure 2.17** shows an optical arrangement in which light is projected onto a diffusing screen, rather than directly onto the object being viewed. This set-up effectively avoids any problems due to glinting and produces soft shadows.

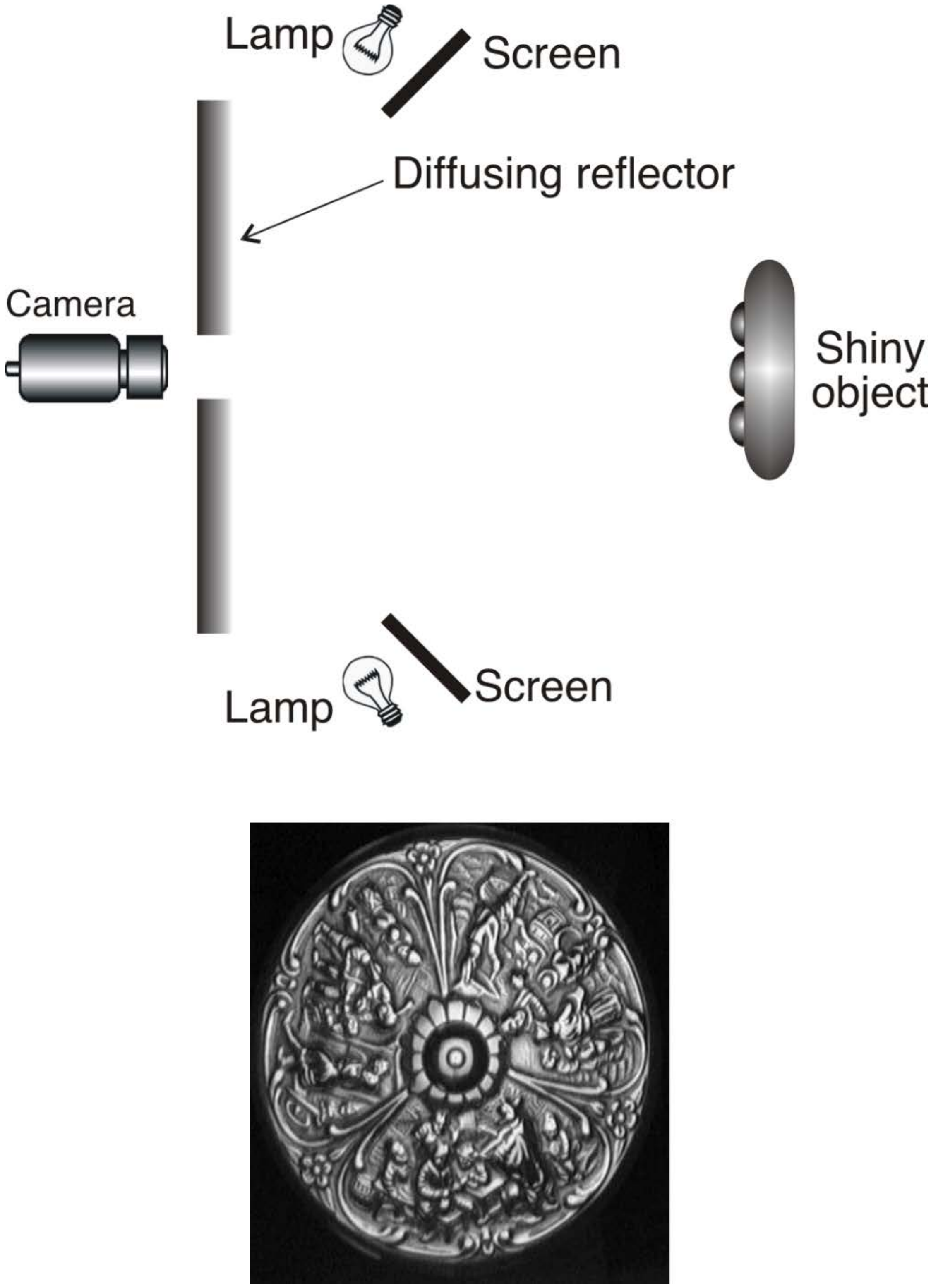


Figure 2.17 *Diffusing reflector effectively acting as the light source for a shiny object. (Top) Optical layout. (Bottom) Sample image, embossed silver tobacco box.*

Dark-field illumination

Various dark-field lighting-viewing arrangements are possible and can be used to to examine glass-ware, clear plastic objects, transparent sheeting, smoke and aerosols. **(Figure 2.18)** Grazing illumination uses light projected at a very low angle to highlight the edges of embossing, while flat surfaces appear dark. Dark-field illumination is also used widely in microscopy. A brick-shaped object can be illuminated in such a way that its top surface is bright, while the back ground is dark. This is useful as an alternative to back-illumination for producing high-contrast images.

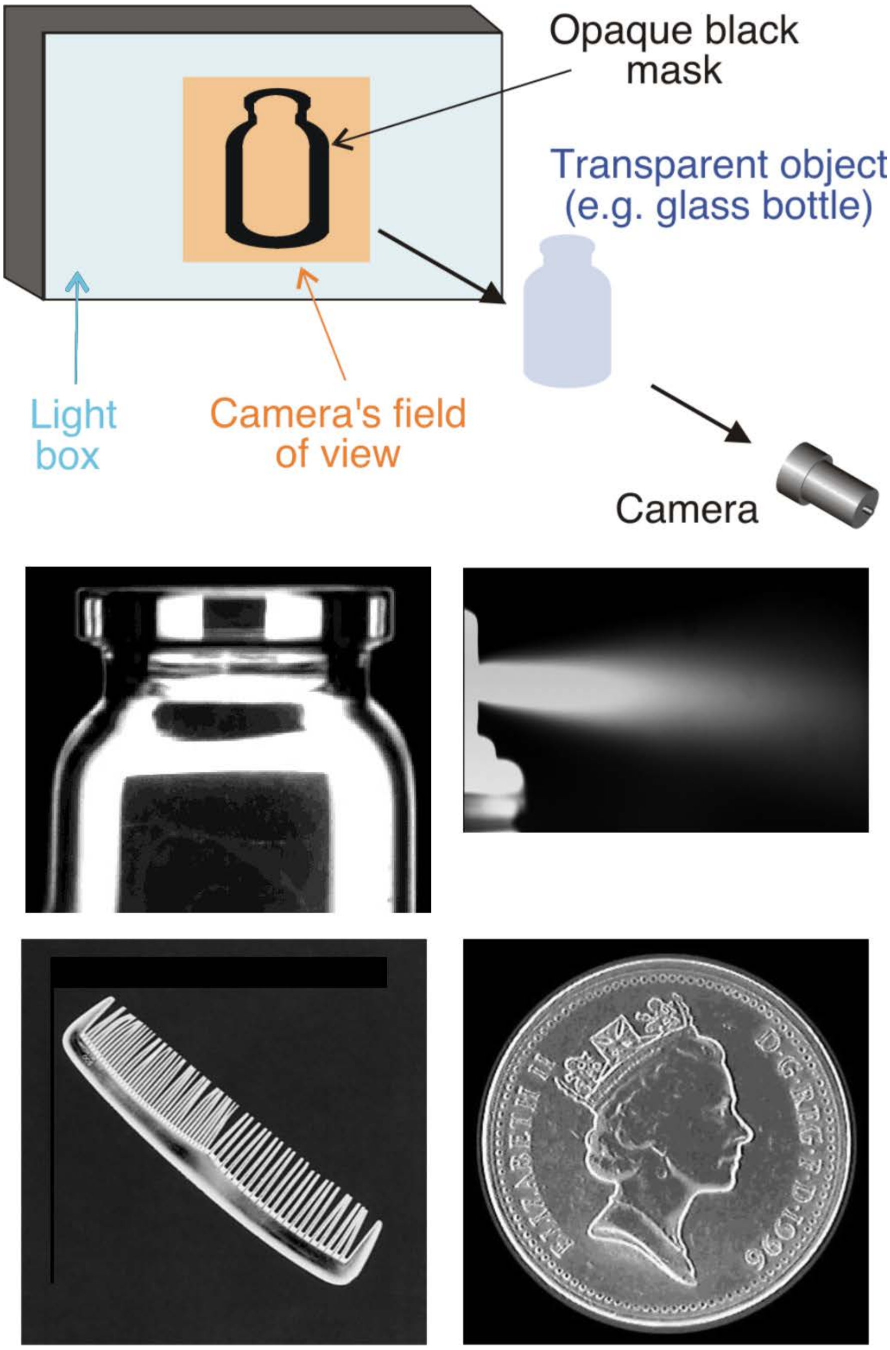


Figure 2.18 Dark field illumination, three variations. (Top) Optical layout for a bottle in fixed position. (Centre-left) Glass vial, illuminated as in [T]. (Centre-right) Aerosol spray cone. Light was projected vertically downwards onto the spray. (Bottom-left) Translucent plastic comb. (Bottom-right) "Silver" coin. Light was projected towards the centre of the coin, at a very low angle, from all directions. This is called Grazing Illumination.

Omni-directional lighting

Shadows and serious effects of glinting can both be avoided by shining light onto the object from all directions. **Figure 2.19** shows how this can be done. Notice that none of the lamps actually casts its light directly onto the object but instead illuminates the inside of the dome.

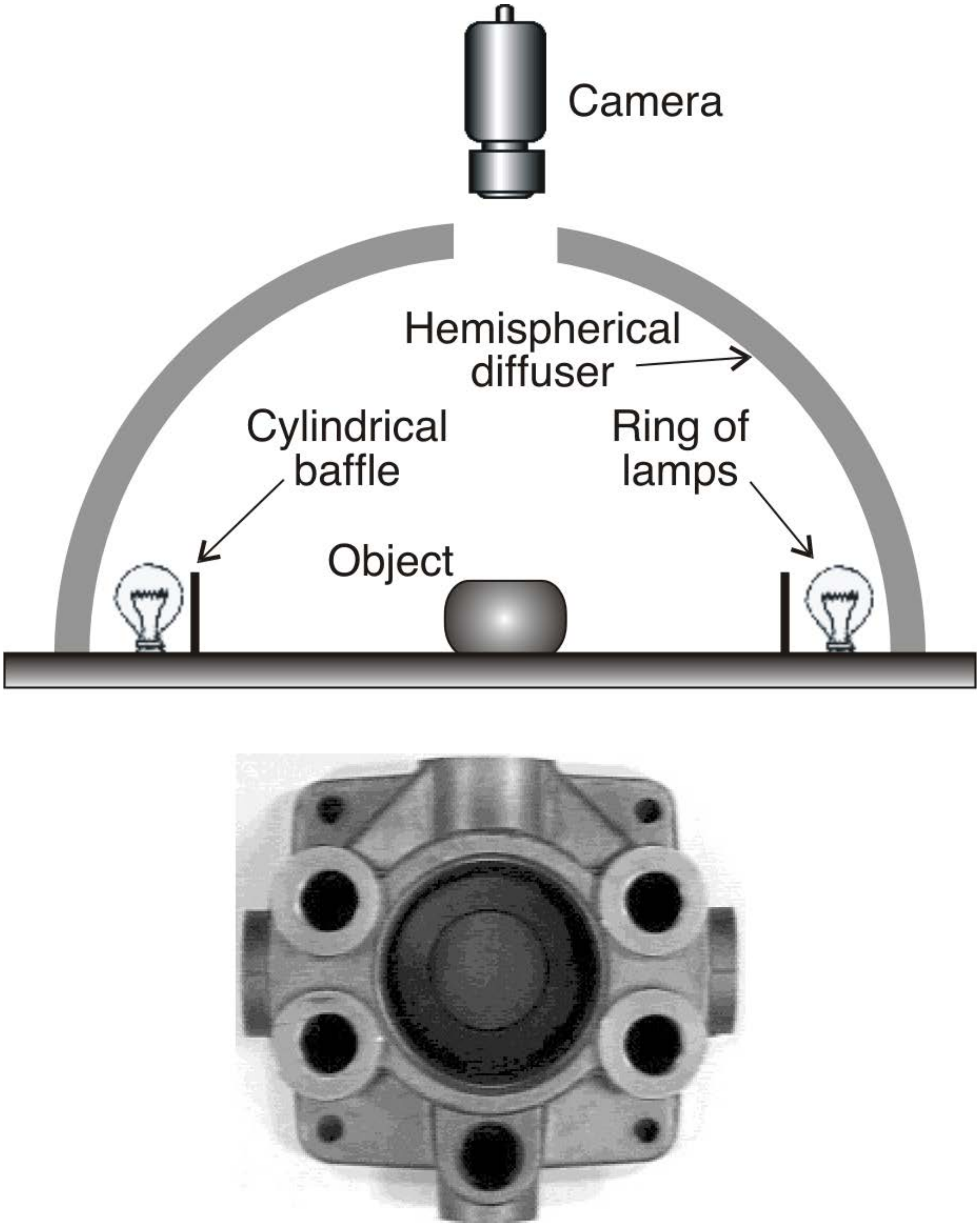


Figure 2.19 Omni-directional lighting. (Also called a "Cloudy Day" Illuminator.) (Top) Optical set-up. (cross-section) (Bottom) Sample image: light-grey alloy casting, about 120x120 mm^2. Step-like features (82 mm high) do not cast significant shadows. The large circular well is 78 mm deep with a diameter of 56 mm.

Structured Light (Triangulation)

This provides an ingenious way to measure the height of a surface. A very thin fan-shaped sheet of light is created, usually by expanding the beam from a laser, and is projected onto the widget surface. The camera is offset from the light source so that it "sees" a bent stripe of light. (Figure 2.20) From this, it is a simple matter to calculate the height of the surface along one vertical "slice" through the widget. By repeating this, many times, with the object shifted slightly at each step, a representation of the 3-dimensional shape of the widget's top surface can be obtained. In a so-called range map *(or* depth map*)* created in this way, surface height is indicated by intensity. Notice that this lighting-viewing method ignores the tone and colour of the surface being viewed.

This optical ranging technique is used to assist navigation in autonomously guided vehicles and to detect pot-holes in roads. Specialist laser scanners, based on the general principle outlined in **Figure 2.20**, are used to obtain 3D shape information about archeological sites, including whole buildings, caves and sewers. [Lidar (laser-based "radar") is often preferred for large structures.] The same scanning technique finds application in dress-making and tailoring, to measure body shape. It is also used, in conjunction with 3D printing, to make prostheses after limb amputation, as well as in other areas of cosmetic and restorative surgery. For example, it is ideally suited to provide precise shape data, prior to breast enhancement surgery. One of the earliest applications was to create a sculpted bust so that the rich and vain can live for ever, in the form of a 3D effigy.

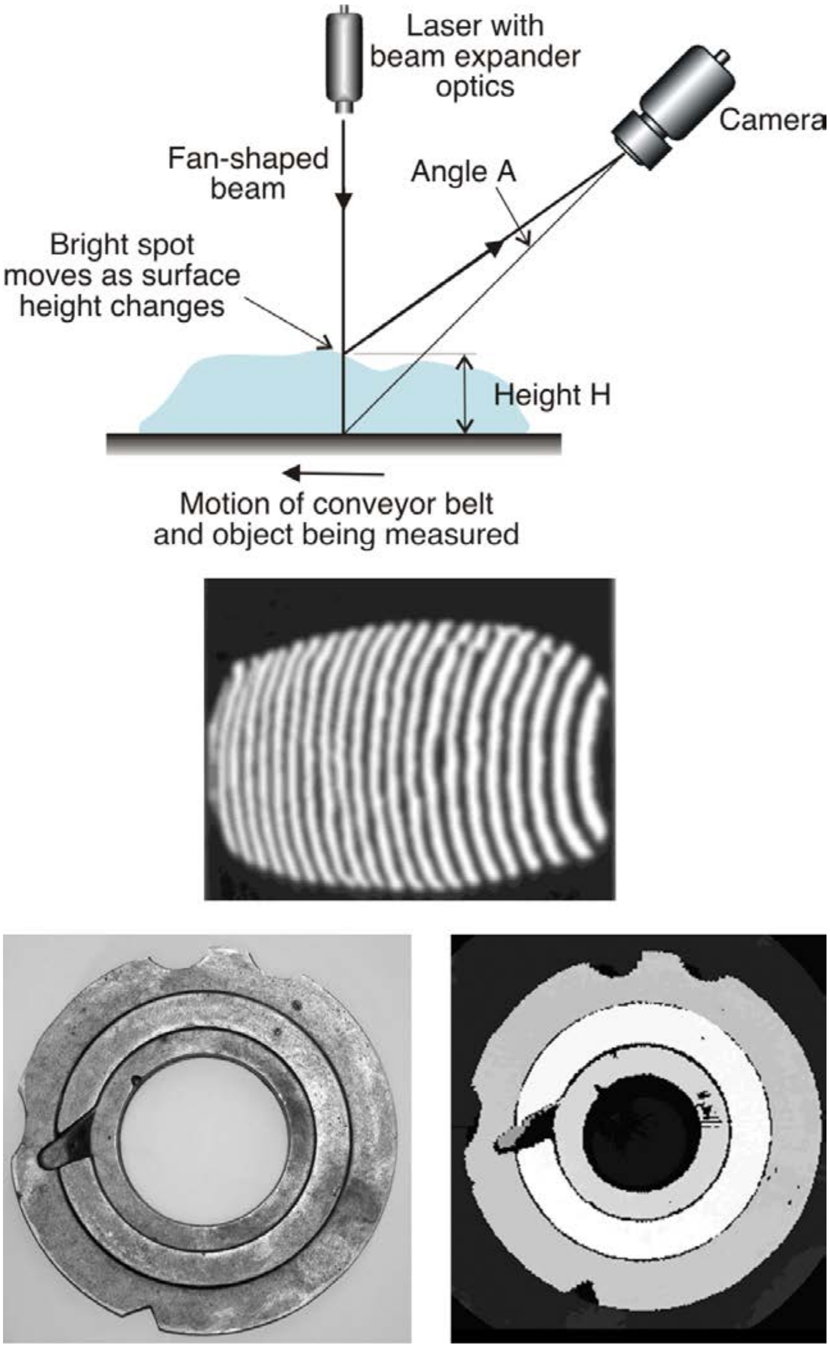
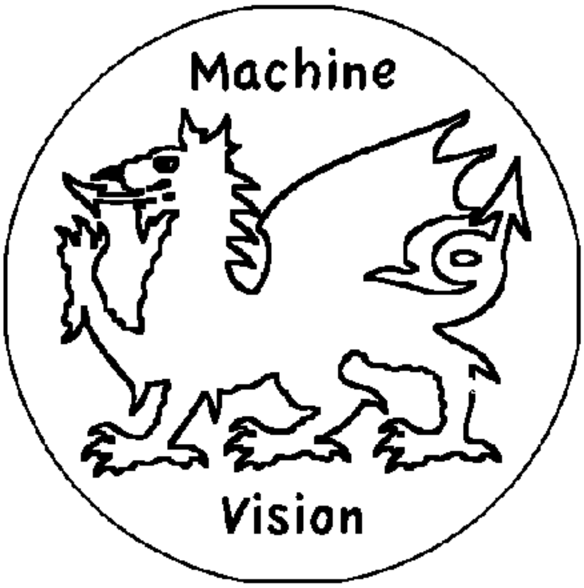


Figure 2.20 Measuring surface height with structured light. (a) Optical arrangement. The fan-shaped beam is viewed edge-on. The camera "sees" a bright wavy stripe. A simple formula relates the surface height (H) to the angle A. (Centre) Multiple light stripes on a bread roll. (Bottom-left) Alloy component for an automobile clutch. (Bottom-right) Range map, built from multiple light stripes.

It is easy to think of many potential applications in industry, particularly in food manufacture: measuring the shape of such ill-formed items such as loaves, cakes, dollops of dough, cream, etc. is important in a modern high-volume bakery. It is also used to measure the shape of solder joints, aerodynamic/hydrodynamic surfaces, car body panels, seating, custom protective clothing, etc.

A small digression: In many early publications the positions of the camera and laser were reversed in diagrams like Figure 2.20. This makes a meaningful range map much more difficult to calculate. This happened originally as a mistake, which was repeated by other authors, including myself, for several years after the original publication. One of my students (Dr. Simon Cotter) discovered this error when a range map derived from a tray of eggs showed the eggs all leaning to one side. Progress does not always follow a straight path!



Blank page

Remarks

In a reference book on Machine Vision published recently, I was able to list over 130 different lighting-viewing methods. These can be used singly, or in various combinations, creating a large number of possible ways to view an object. The small sample listed here includes the most popular. Notice that the arrangement is Figure 2.13 is able to reproduce the same effects as outlined in Figures 2.14 - 2.19.

Specialist optical devices with integrated lighting and viewing are are able to produce some spectacular effects. Four examples are shown in **Figure 2.21**. A well-informed engineer can often exploit advances like this to good effect, making the vision system faster, cheaper and more reliable. Clearly, we cannot illustrate more than a tiny fraction of these.

Figure 2.21 Specialist lenses can sometimes make inspection much easier. These images were all obtained using a hypercentric lens. (Top-left) Cube, gaming die. Five of the six faces are visible simultaneously. (Top-right) Internal view of a metal cylinder with six side ports. (Bottom-left) Cylindrical adhesive stick container. Notice that lettering all round the outside surface of the cylinder is visible. (Bottom-right) Internal view of a car oil filter, viewed through a small hole in the top. (These images were kindly supplied by Spencer Luster, Light Works, LLC, Toledo, Ohio, USA.)

Cameras

The development of cheap, high-performance solid-state cameras has lead to significant improvements, in recent years, in road traffic monitoring, domestic, commercial and military surveillance, burglar alarms, camcorders, webcams, digital photography, inspecting drains, dentistry, medical endoscopy, etc. Modern cameras have also benefitted manufacturing industry and they are central to most industrial Machine Vision systems. In some specialist applications, old-fashioned tube cameras and other exotic image sensors are used instead but CCD and CMOS cameras are used for a large majority of industrial applications and will be the main focus of our attention here.

It is not necessary to dwell on the details of the technology of modern cameras but we do need to understand the basic principles. A lens focusses an image onto the camera's light-sensitive surface ("retina"). **(Figure 2.22)** This consists of a regular array of very small photo-detectors (light detectors), formed on the flat surface of a small wafer of semi-conductor material. When light falls on one of these photo-detectors, a small but measurable electric charge is created. The magnitude of this charge increases with the duration and intensity of the illumination.

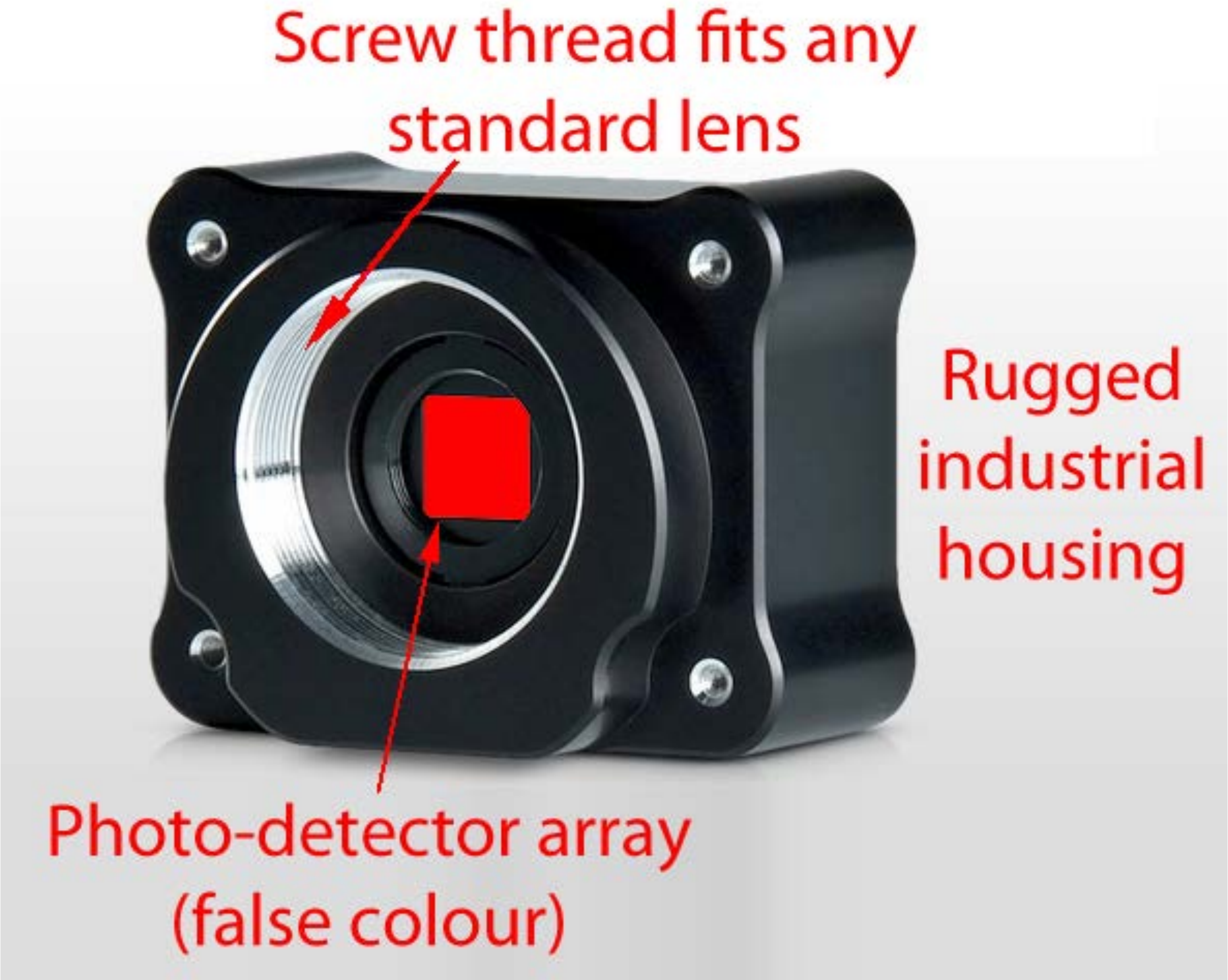
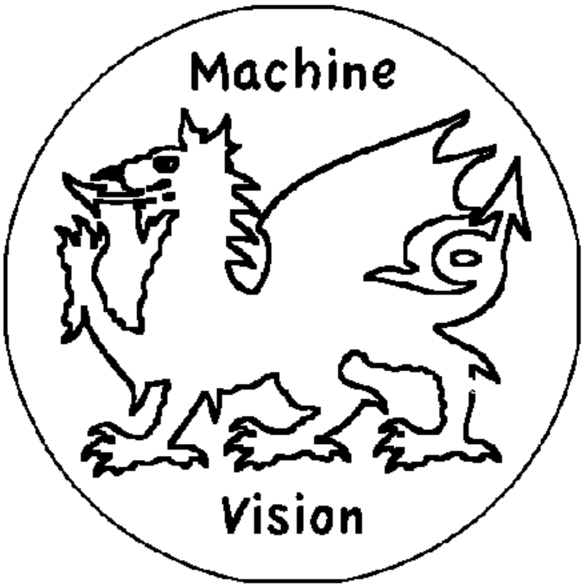


Figure 2.22 Semi-rugged industrial camera with a rectangular photo-detector array ("retina") and a standard screw thread. The latter can accommodate a variety of lenses, including macro, fish-eye, telephoto, endoscope, microscope, telescope, fibrescope, etc.

It may be helpful to think of this charge-creation process as equivalent to spraying drops of water onto a flat surface. Clearly, the amount of water collected increases with the number of drops added each second and the length of time the spray is turned. on. The latter is called the integration interval and the spray density *(measured in drops/second)* is analogous to the light intensity*.* A narrow beam of light generates a local concentration of charge, perhaps within a single photo-detector site, so the image of a scene forms an equivalent charge pattern in the silicon.

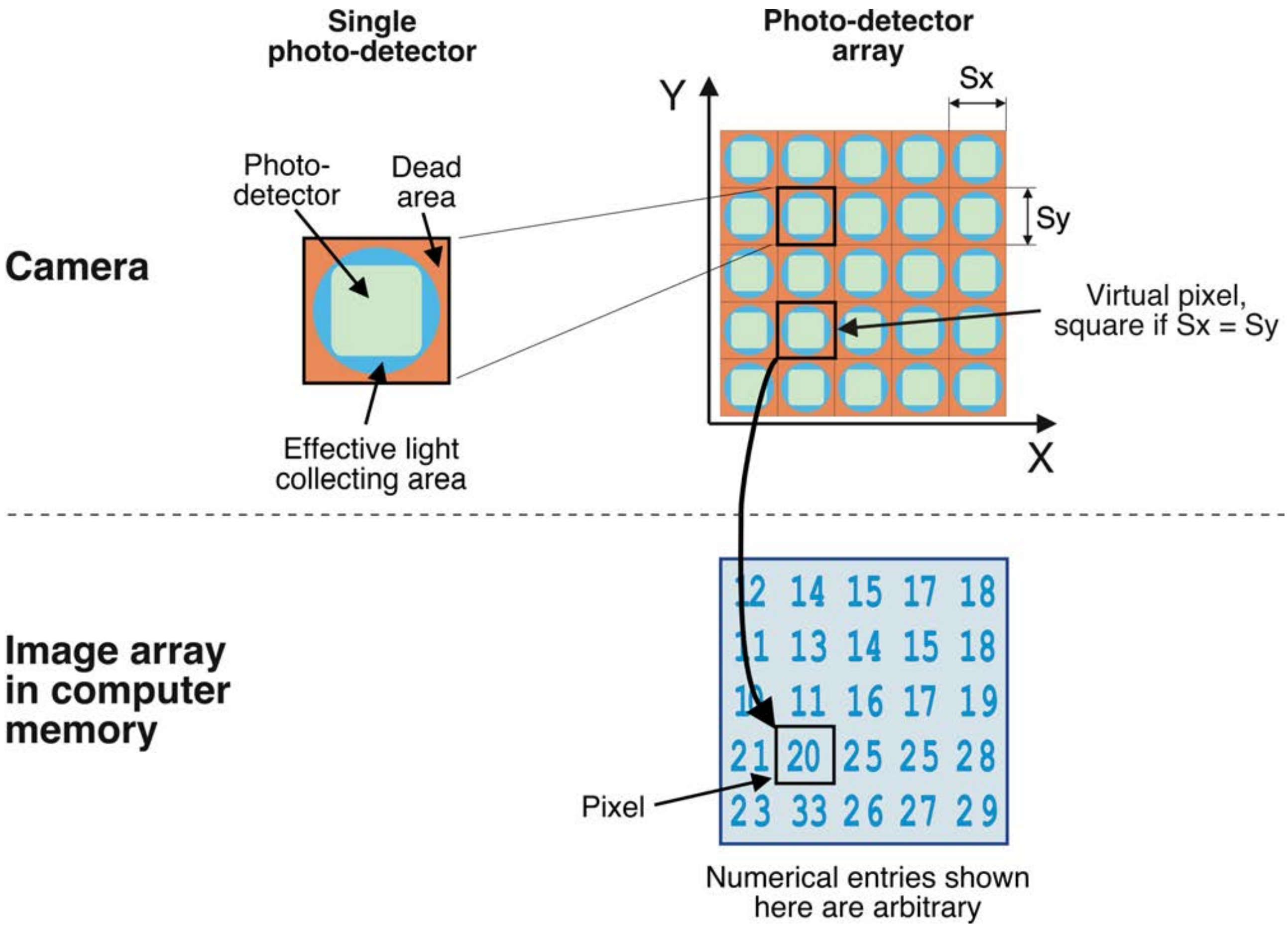
At the end of the integration interval, an electronic circuit built into the same semi-conductor wafer senses each photo-detector site, in turn, and detects how much charge has been collected. It converts this into an electrical voltage (analogous to water pressure). Hence, a pattern of light is converted into an equivalent pattern of electrical charge, which is finally converted into a sequence of voltage changes. This constitutes the camera output. The scanning sequence used to sample the photo-generated charge pattern is called a raster scan and resembles the gaze path followed by the eyes when reading.



Blank page

Capturing Image Data

The camera output is a so-called analogue signal. That is, it can vary smoothly between defined limits. A high output voltage represents a bright point in the light pattern projected onto the camera's retina and, of course, a low voltage denotes a dark point. Unfortunately, computers do not accept analogue signals, so a circuit, called and an analogue-to-digital converter, is needed. Nowadays, this is built into a camera, justifying the use of the term digital camera. The signal flowing from the camera into the computer, is in the form of a stream of numbers, each represented in digital form. If we think of the raster scan as "unfolding" a 2-dimensional charge pattern, the reverse process is needed, to create a 2-dimensional table (called an array) of numbers held inside the computer memory. This is called an image array and is the starting point for computer processing of pictures. **(Figure 2.23)** We will return to this in the next chapter.

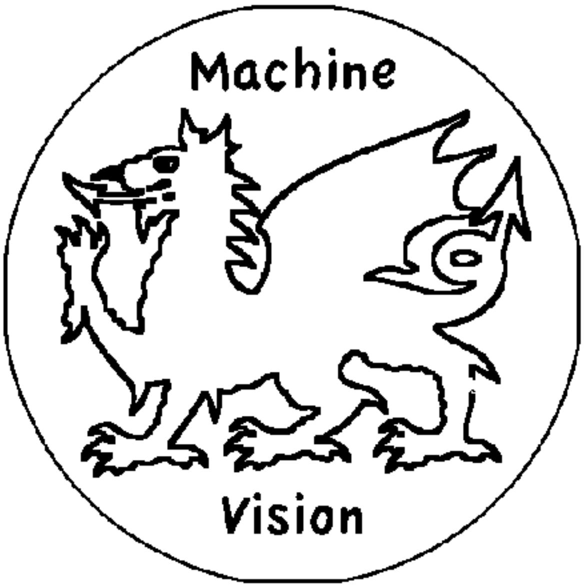


12	14	15	17	18
11	13	14	15	18
10	11	16	17	19
21	20	25	25	28
23	33	26	27	29

Figure 2.23 Relating the sensor array of a camera to a monochrome digital image. The spacing of photo-detectors along the X and Y axes is normally the same (Sx = Sy). This ensures that the virtual pixels on the digital mage are square. Many cameras are fitted with a tiny lens over each photodetector. This improves image sharpness and makes the camera more sensitive, by effectively increasing the light collecting area.

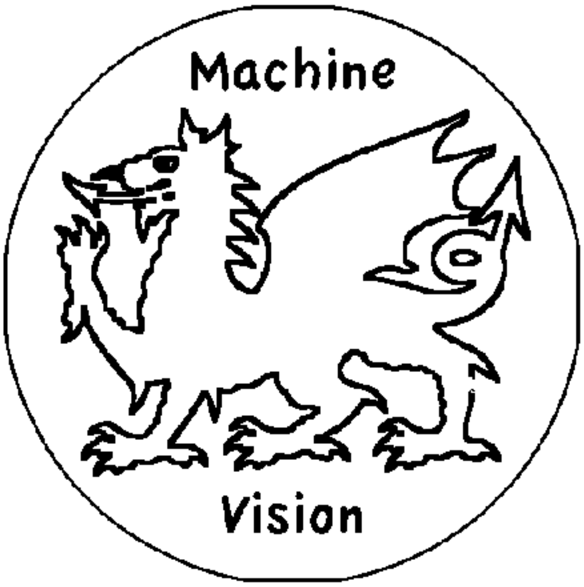
Pixels

Many people first become aware of the word pixel when they buy a digital camera. As far as the customer and manufacturer are concerned, the greater the number of megapixels that a camera boasts, the better. Pixels are often equated with photo-detectors but this is a mistake! A photo-detector has a physical existence on a camera's retina, whereas a pixel is, strictly speaking, a *position* in an image array, where a number (or several numbers) is held. A manufacturer might, for example, produce a camera with, say, 1000*1000 separate photo-detectors but claim it can resolve 2000*2000 pixels. (This is possible if "extra" pixels are generated, by estimating what light level a hypothetical photo-detector, located mid-way between two existing photo-detectors, would detect. This involves a mathematical process called interpolation.) Later, we will see that colour cameras might combine four photo-detector outputs to form a single pixel. For the moment, we will, for the sake of conceptual convenience, assume that every pixel is uniquely and exclusively associated with one photo-detector.



Blank page

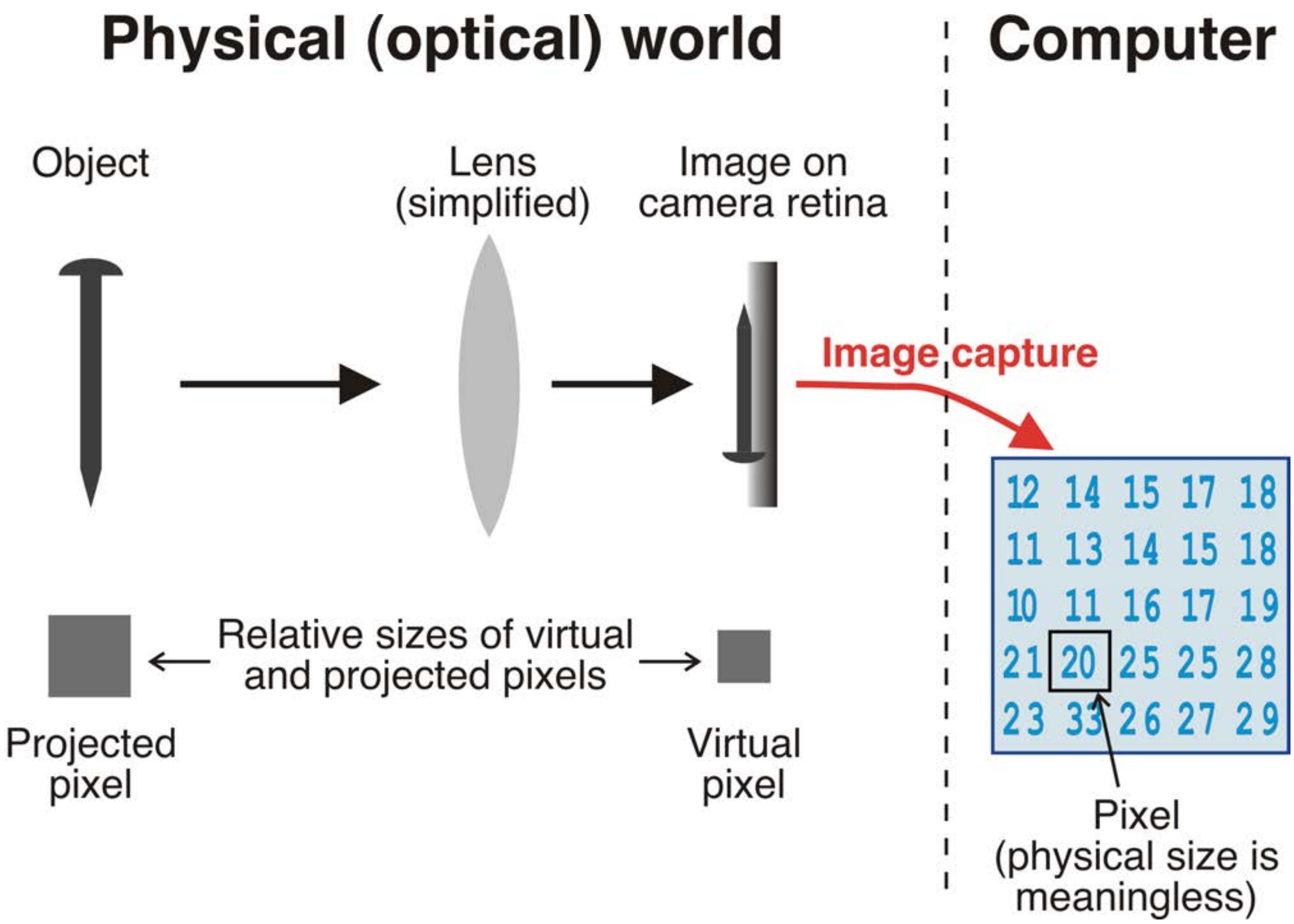
So, a pixel is a position, or slot, in an array of numbers and has no physical existence. Despite this, we often discuss the size and shape of a pixel, as if it did exist in reality. A pixel has a value (i.e. a number) that attempts to describe the average light level falling on its associated photo-detector over the integration interval. We are only interested in patterns of light, so the size and shape, of the photo-detectors is unimportant, provided they are all the same. With this in mind, we often find it convenient, to *think* of a pixel as a rectangular region on the retina. (See the squares with black borders in Figure 2.23) Let us call this a virtual pixel. The shape of the virtual pixels is important. If they are not, the size of an object, measured by the number of pixels it covers, appears to vary as it is rotated. In the early days of Machine Vision, manufacturers did not fully appreciate the inconvenience of non-square virtual pixels but vision equipment is nearly always designed nowadays with square pixels, to avoid these problems.



Blank page

Suppose you were to stand at the position of the object being viewed and look into the camera lens. The virtual pixels would appear to be enlarged by the lens. This gives us another idea: projected pixels. **(Figure 2.24)** The size of a projected pixel is 1/M times that of the virtual pixel, where M is the magnification factor of the lens. The size of projected pixels are important because this defines the size of the smallest feature that the camera-lens combination can resolve. (It is about four times the size of the projected pixels, for reasons that will be explained in Chapter 3.)

Sometimes, equipment manufacturers claim to be able to make measurements with sub-pixel precision. Is this possible? It is in certain situations. For example, to measure the diameter of a wire with "sub-pixel precision"; simply take lots of diameter measurements and average them. The result will be a more precise estimate than a single measurement could yield.



12	14	15	17	18
11	13	14	15	18
10	11	16	17	19
21	20	25	25	28
23	33	26	27	29

Figure 2.24 Pixels, virtual pixels and projected pixels. Inside the computer, an image is usually represented as a table (array) of numbers indicating the brightness for each pixel. The magnification of the lens (M) is the ratio of the image size to object size. (Here, M is less than 1.) Hence, the size of the virtual pixels is M times that of the projected

Colour Cameras

Thus far, we have described how a monochrome camera works. It yields one measurement of intensity for every pixel. Most colour cameras take three such measurements for each pixel. These are usually referred to as R, G and B. (These signals do not measure the red, green and blue components exactly but the approximation is good enough for our present discussion.) **Figure 2.25[T]** shows in diagrammatic form how a 3-sensor camera can be constructed. This arrangement is well suited to large sensors but another arrangement better suits modern photo-detector arrays. Each element in the array is overlaid with a tiny colour filter. In the so-called, Bayer mask, a group of four photo-detectors derives the RGB outputs for a single pixel. **(Figure 2.26[B])** The photo-detectors covered by the two green filters are averaged to generate the G signal. The filter mask is created during manufacture and does not need the very careful alignment that the arrangement in Figure 2.25[T] requires. The signal from the R-channel effectively defines a monochrome digital image, which we will call the R-image. The G-image and B-image are defined in the same way. A colour image therefore has three component monochrome images, which can be separated easily and then processed individually like any other monochrome image.

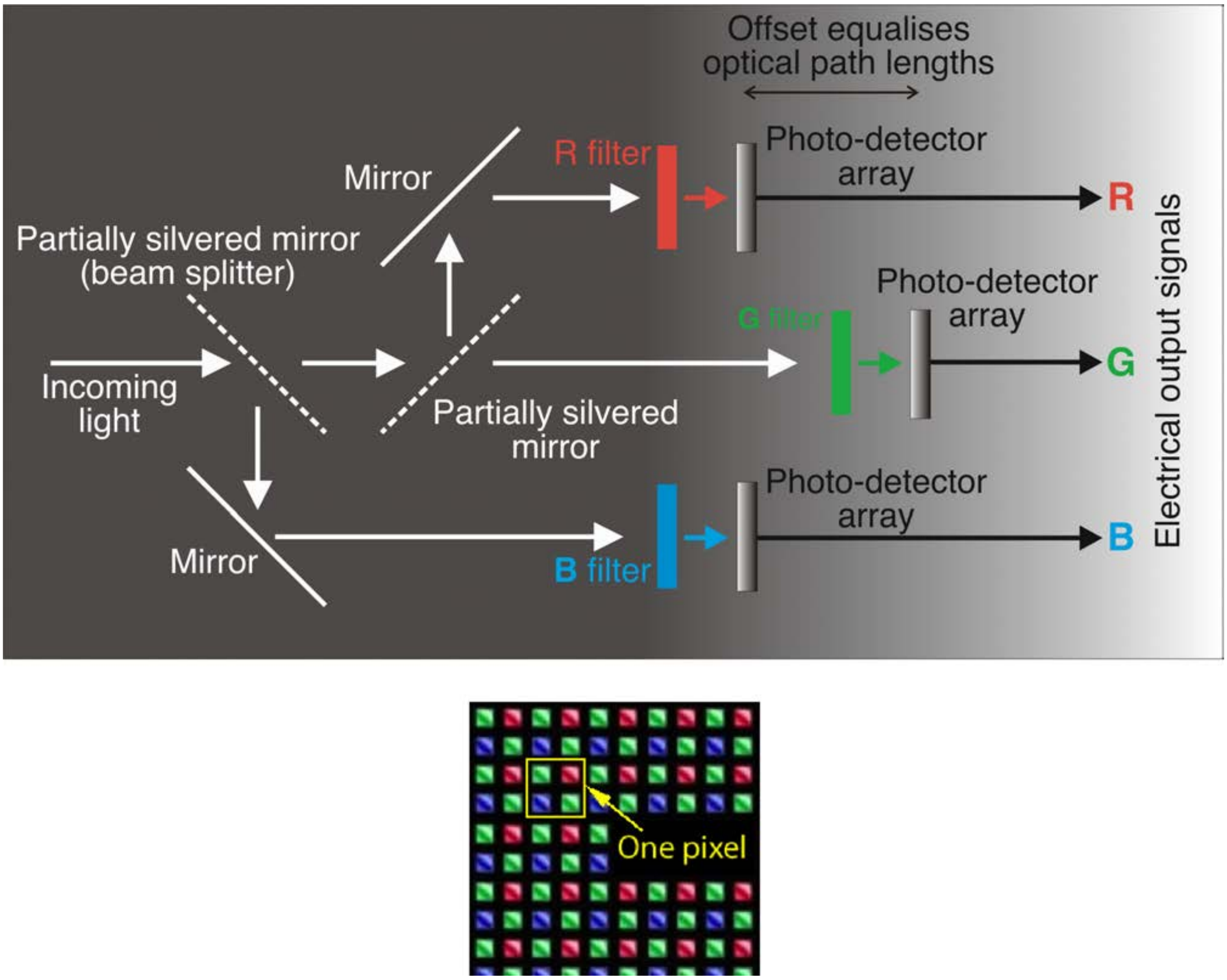


Figure 2.25 Colour camera. (Top) Diagrammatic representation of a 3-sensor camera. The left of this diagram (dark background) represents optical signal flow, while the photo-detector arrays generate electrical signals. (Bottom) Arrangement of the Bayer filter mask for a single-sensor colour camera. One photo-detector is located behind each colour patch. Four photo-detectors on the same chip are needed to define the RGB values for a single pixel. The signals from the two green photo-detectors are averaged.

Line-scan Cameras

Ordinary, cameras, such as those used for photography, television and security applications, have their photo-detectors arranged to form a 2-dimensional array. A line-scan camera has all of its photo-detectors arranged in a single straight line. **(Figure 2.26)** It can, therefore, detect only linear light patterns. This may seem rather restrictive but, in fact, it is very convenient in certain situations.

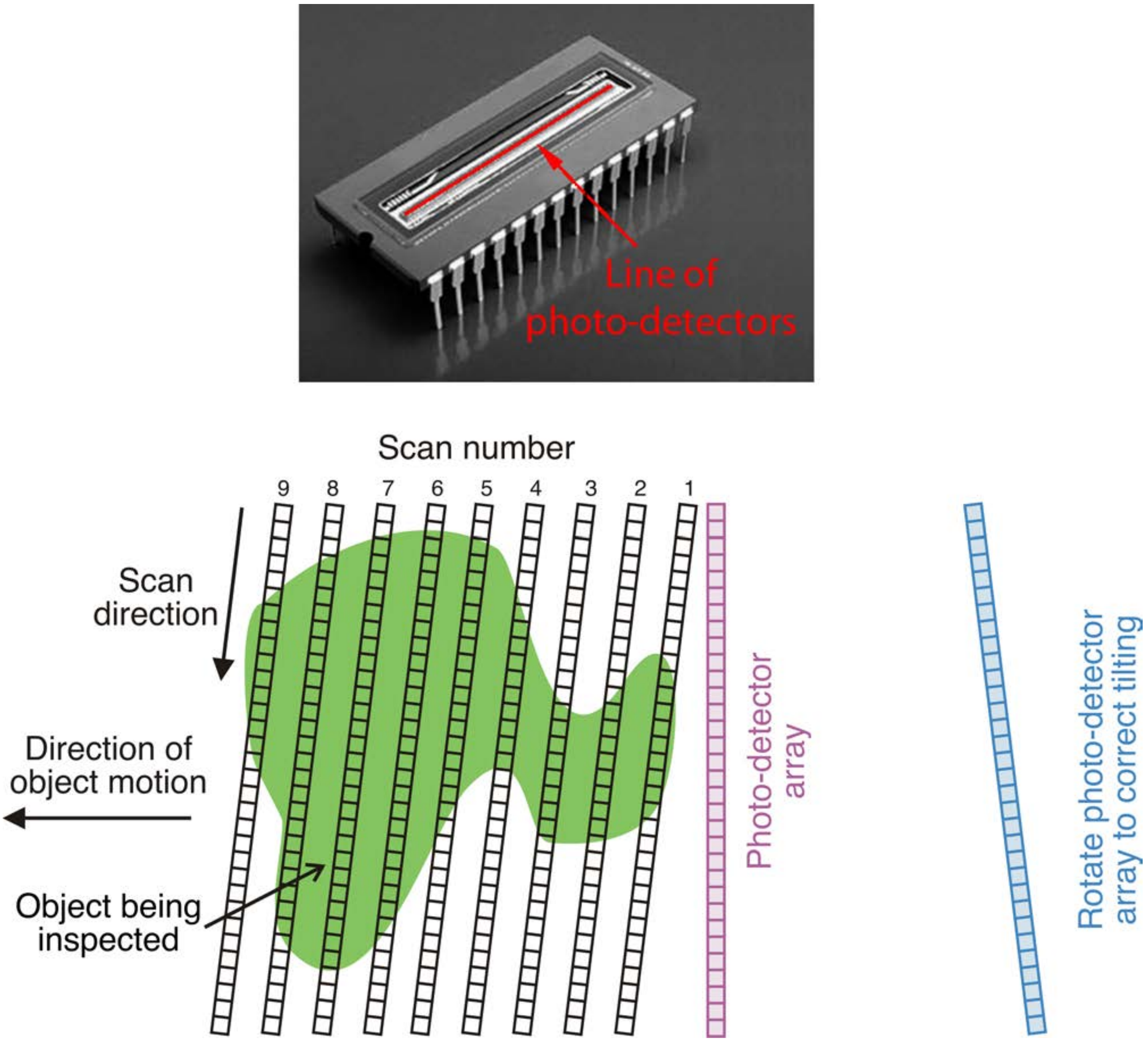


Figure 2.26 Line-scan camera. (Top) There may be as many as 4096 photo-detectors arranged along a straight line, on a single camera chip. (Bottom) Repeated scanning allows a 2-dimensional image to be constructed. This introduces trapezoidal warping in the composite image. To compensate for this, the line-scan camera should be rotated slightly

It is possible to build up a 2-dimensional image of an object as it moves past a linear array that is repeatedly scanned. Hence, a line-scan camera is often used in conjunction with a continuously moving factory conveyor belt, or to inspect indefinitely long webs (paper, steel sheet, etc), or products, such as extruded confectionary bars, wire, pipes, carpets, cloth, etc. In fact, a line-scan camera is more convenient than a regular array camera for these types of product.

Figure 2.27 shows an image obtained by a line-scan device from an object rotating continuously on a turntable. Figure 2.27[B] shows a bit of software magic: the line-scan image has been warped, to reconstruct a normal 2-dimensional view. (The image has been "bent round in a circle", so that the top and bottom rows are adjacent.) Notice however, that it would be extremely difficult to illuminate the object in a way that would recreate this image using a normal 2-dimensional sensor array. In other words, a human observer would never normally see an image like this, except by computer processing.

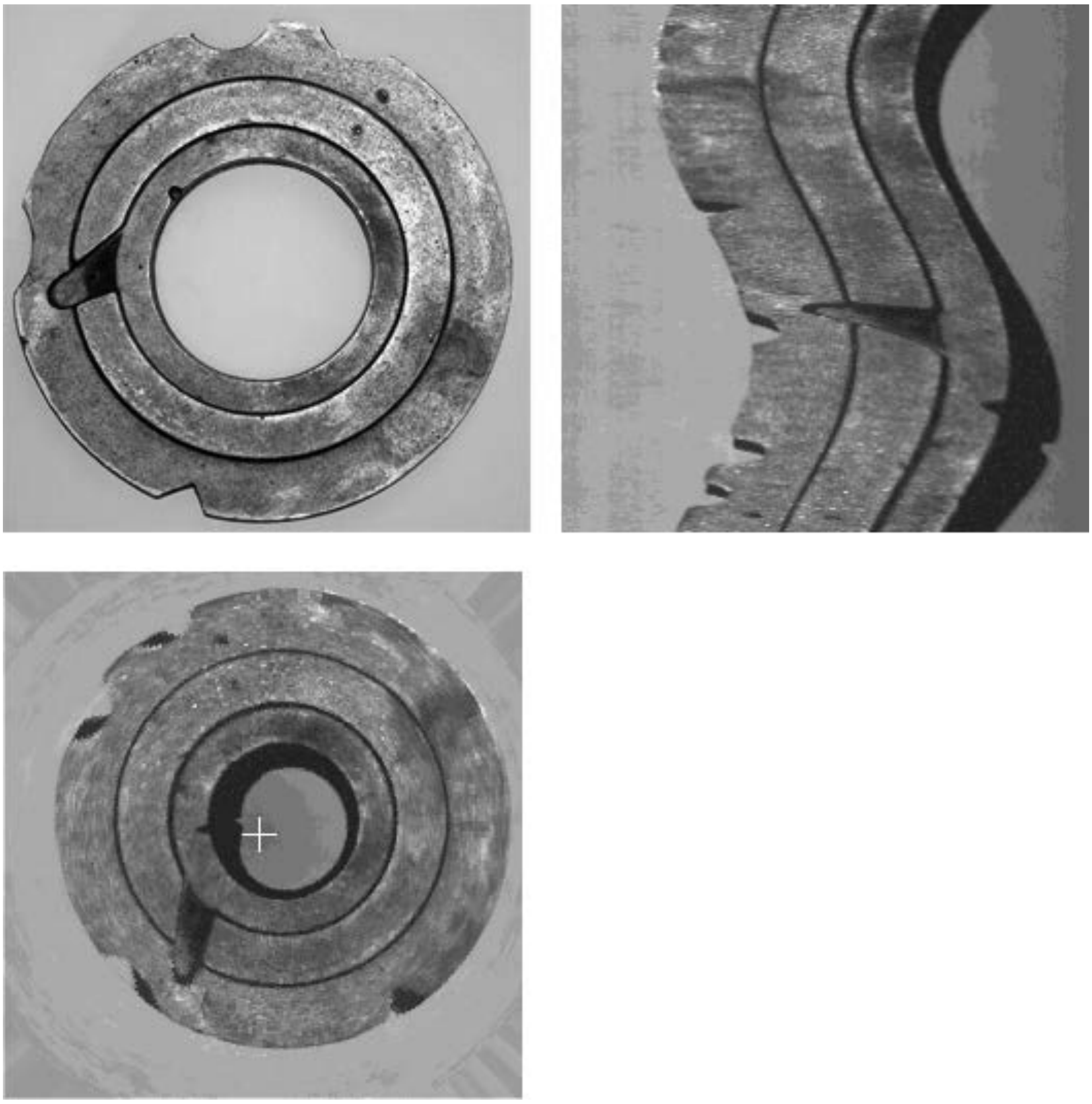
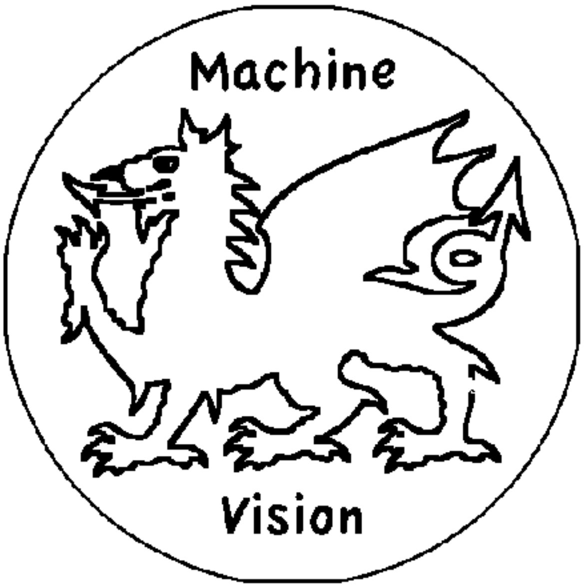


Figure 2.27 A line-scan camera views a circular object on a continuously rotating turntable. (Top-left) Normal view. [Cast alloy, part of an automobile clutch] (Top-right) 2-dimensional composite image, constructed by repeatedly sampling the output of the line-scan camera as the object is rotated. If this toroidal component were concentric with the turntable, this image would would consist of a set of straight vertical stripes. (Bottom) Image reconstructed from [TR] by software, which effectively "bent" the image to form a circle The cross is the centre of the image and therefore corresponds to the centre of the turntable.

Image Processing

The fact that a computer can store and process pictures is now very familiar. Anyone who has ever used an image editing program such as Photoshop®, the touch-up software provided with a digital camera, or participated in a web-based video conference (e.g. Zoom) will be aware of some of the possibilities. Television and cinema drama productions frequently incorporate computer enhanced imagery. (This is distinct from Computer Generated Imagery, CGI, which has given us the chance to see dinosaurs as if they were alive today.) The Internet provides ready access to a huge range of short movie clips. Camcorders and DVD players both have built-in computers that store and process images. This is public knowledge now but even recently (1990s) few people even suspected that these wonders would be possible. Although image processing by computer is commonplace now, few people appreciate how this is achieved. We will explain some of the mysteries of image processing in later chapters. What will be demonstrated is probably quite different from anything that you have imagined before; there are no spectacular Hollywood-style effects but the results have a greater practical value, by improving product safety and reducing costs.



Blank page

Systems

So far in this chapter, we have examined parts of a Machine Vision system but have not yet discussed its overall structure. **Figure 2.28** is one of the most important diagrams in this book as it emphasises that a vision system brings together lots of different technical subjects. Although they are very different, these parts must all be integrated to work together smoothly and harmoniously. However, it must do more than that, since an MVS is part of a bigger entity: a factory. The latter is, in turn, part of an even bigger item: a company. Let us now consider in detail how a vision system works in this context. The following section can be omitted without serious loss.

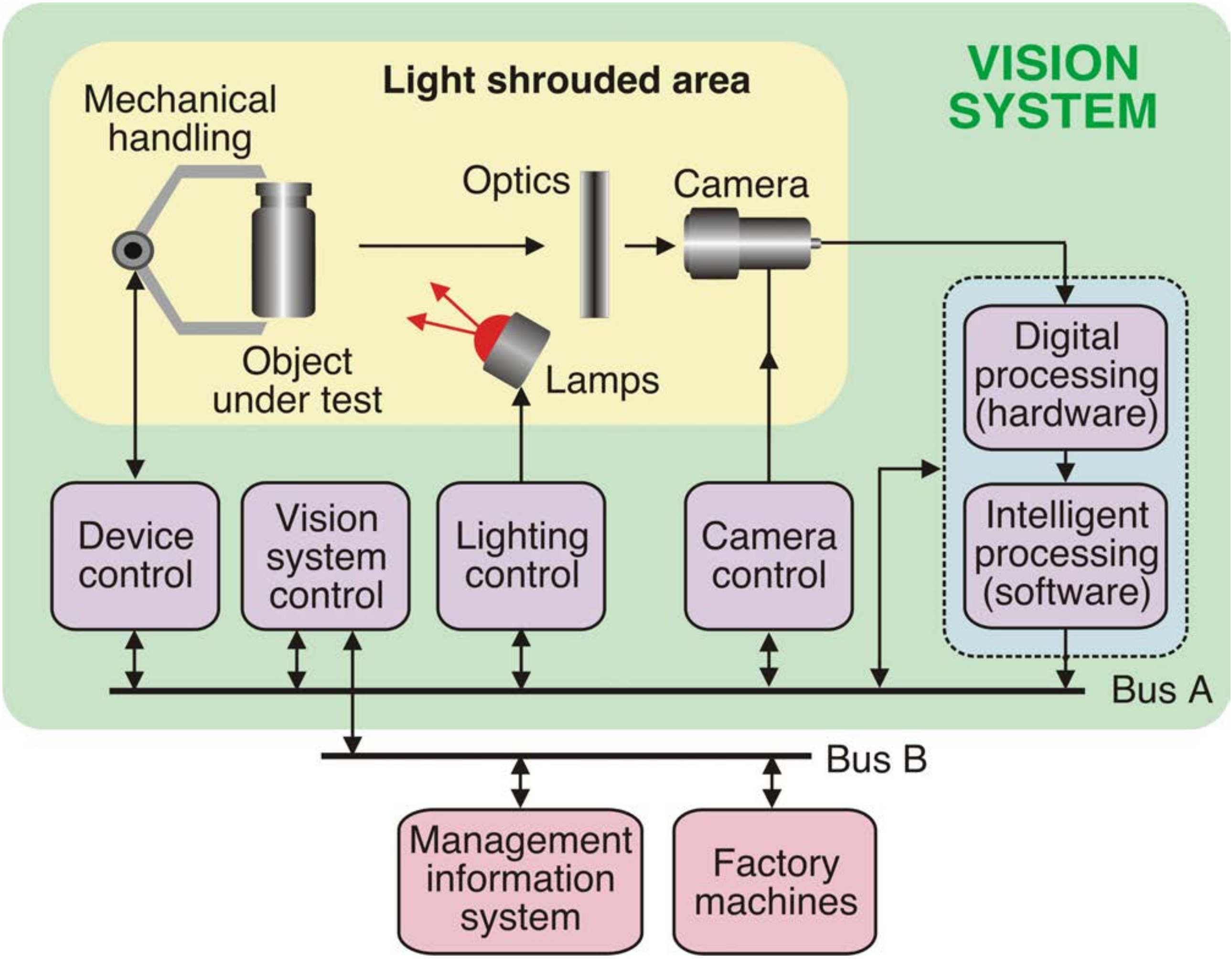


Figure 2.28 Internal organisation of a typical Machine Vision system, integrated into the working environment of a factory and the company-wide information-management system. Buses A and B are data-flow paths. The blue box encloses the image processing modules. Fast image processing is performed within the box labelled "Digital processing (hardware)", while more complex actions are executed in "Intelligent processing (software)". All actions within the vision system are coordinated by "Vision system control".

Automated Visual Inspection System

Refer to Figure 2.28.

The box labelled *"Vision system control"* represents the central controller for the whole vision system. It coordinates the actions of all other parts, by sending and receiving signals via "*Bus A"*. This is a multi-core cable carrying data at high speed between modules connected to it. Thus, the *"Vision system control*" box might issue a command to the box labelled *"Device Control"* ordering it to place the object currently held by the mechanical handling device in the reject bin.

The arrival of a new object for inspection may be signalled in one of two ways:

A signal is sent from *"Factory machines",* via "*Bus B"*, to *"Vision system control"* .

The mechanical handling device has a sensor which detects the presence of the new object. This initiates a signal being sent via *"Device control"* and "*Bus A"* to *"Vision system control"*. This sensor might typically be a photo-cell, that detects when a light beam has been interrupted

"Vision system control" sends a signal to *"Camera Control"* ordering a new image to be captured. *"Vision system control"* might also order *"Lighting control"* to flash a strobe lamp, at the same time.

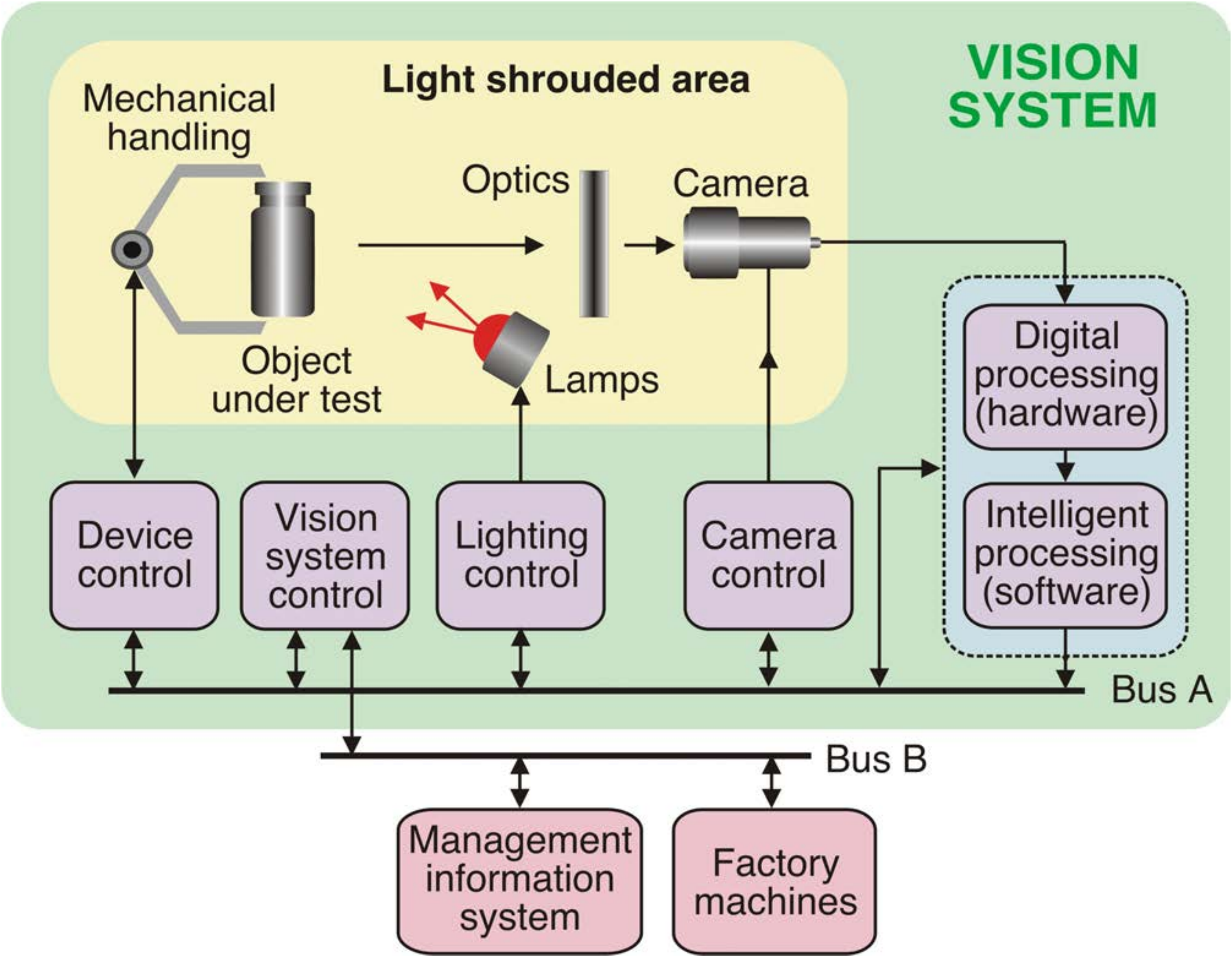


Figure 2.28 (Repeated)

The output from the camera is sent to *"Digital processing (hardware)",* which contains a fast electronic circuit dedicated to processing image data. This unit can typically perform only a limited range of operations but does so extremely rapidly, thereby reducing the load on the *"Intelligent processing (software)"* module.

The module labelled *"Intelligent processing (software)"* is able to perform a wide range of processing, analysis and measurement operations on image data. It eventually produces a simple result (typically *accept/reject* the object), or a series of measurements, such as size, orientation, position, etc.

The results just calculated are sent via "*Bus A*" to the *"Device control"* module for action. This may consist of a very simple operation, such as discarding a faulty object. Another possibility is to use image measurements to guide a multi-axis robot to pick up the object, wherever it is resting, and place it in standard position and orientation

Both the *"Camera control"* and *"Lighting control"* units are likely to be capable of acting autonomously, to optimise image quality. This is important as lamps become dimmer with increasing age.

Adjustments to the camera, lighting, mechanical handling and image processing modules may be required when the factory switches to different types of products. *"Vision system control"* coordinates this.

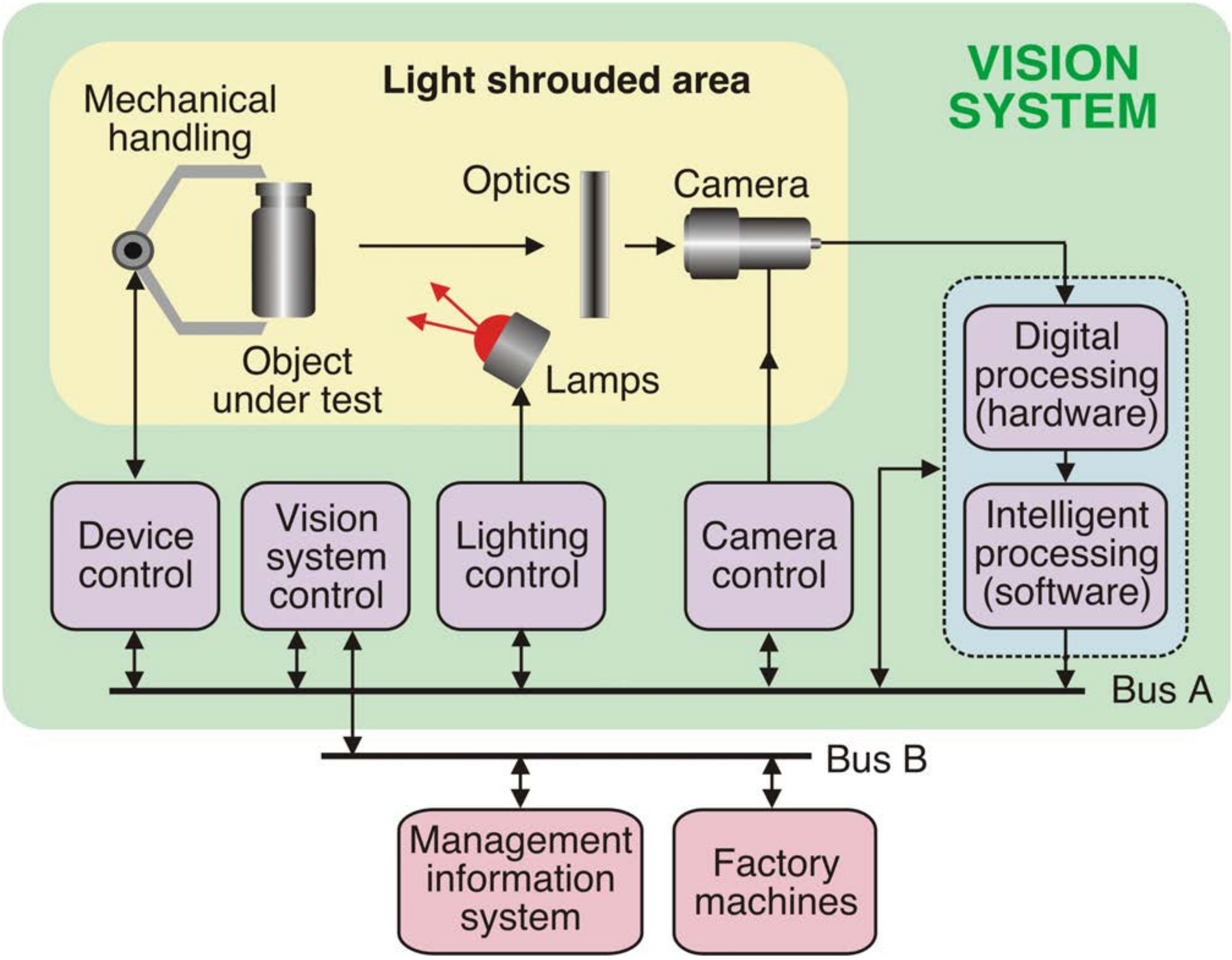
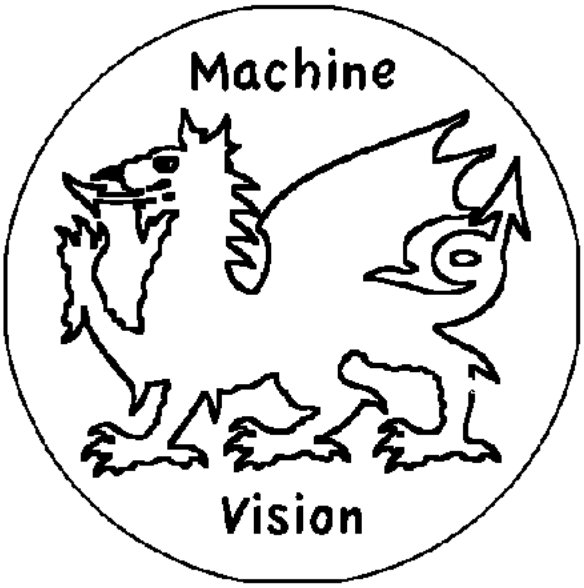


Figure 2.28 (Repeated)

The vision system may control factory machines, such as drills, lathes, grinding, etc. To maintain a clear line of control, only the *"Vision system control"* module is allowed to communicate with *"Factory machines"*, via *Bus B*.

Another useful function that an automated visual inspection system can perform is collecting production statistics. This data is made available to company managers via *"Management information system"*.



Blank page

Systems Thinking

We conclude this chapter by explaining how "systems thinking" can improve the design of a system for inspecting objects on a conveyor belt. There are several points worth noting:

Continuously moving conveyors are often used in factories. Compared to stop-start (indexed) conveyors, they are cheap, energy efficient, reliable and easy to install and maintain, as well as being quiet and simple to control.

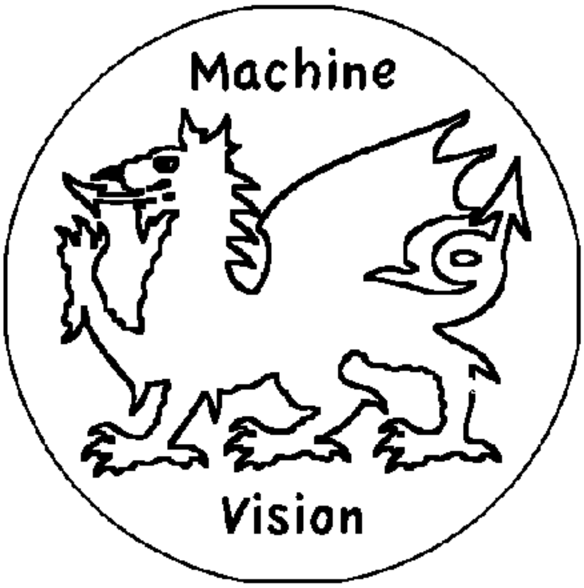
Conveyor belts are commonly found in factories, in which case no special product-handing facilities are needed.

Using an area-scan camera to inspect objects in continuous motion requires either a stroboscope (dangerous), or a camera with an electronic shutter (expensive).

Flashing lights can induce epileptic fits, migraine and make rotating machines appear stationary, so the inspection area must be enclosed to keep strobed lighting in and ambient light out.

A line-scan camera avoids these difficulties. However, it does require that the speed of the conveyor is known and preferably held nearly constant.

When fitted with suitable guide rails, the motion of the conveyor belt can often be exploited to manoeuvre objects into a well-defined position across the conveyor. This applies to some industrial artifacts but, of course, not all.



Blank page

A special kind of computer called a concurrent processor array can be used to inspect objects on a conveyor very efficiently. **(Figure 2.29)** This is quite different from a conventional computer. It has many separate but identical image processing modules (IPs), which together increase the inspection throughput rate (i.e. number of objects examined per minute). All IPs operate an identical program.The time taken by each IP to perform its calculation may be many times longer than the transit time over a distance equal to the spacing between consecutive objects on the conveyor. This can be tolerated because there are many IPs working at the same time, while the conveyor belt holds each object from the moment of image capture to its reaching the accept/reject mechanism.

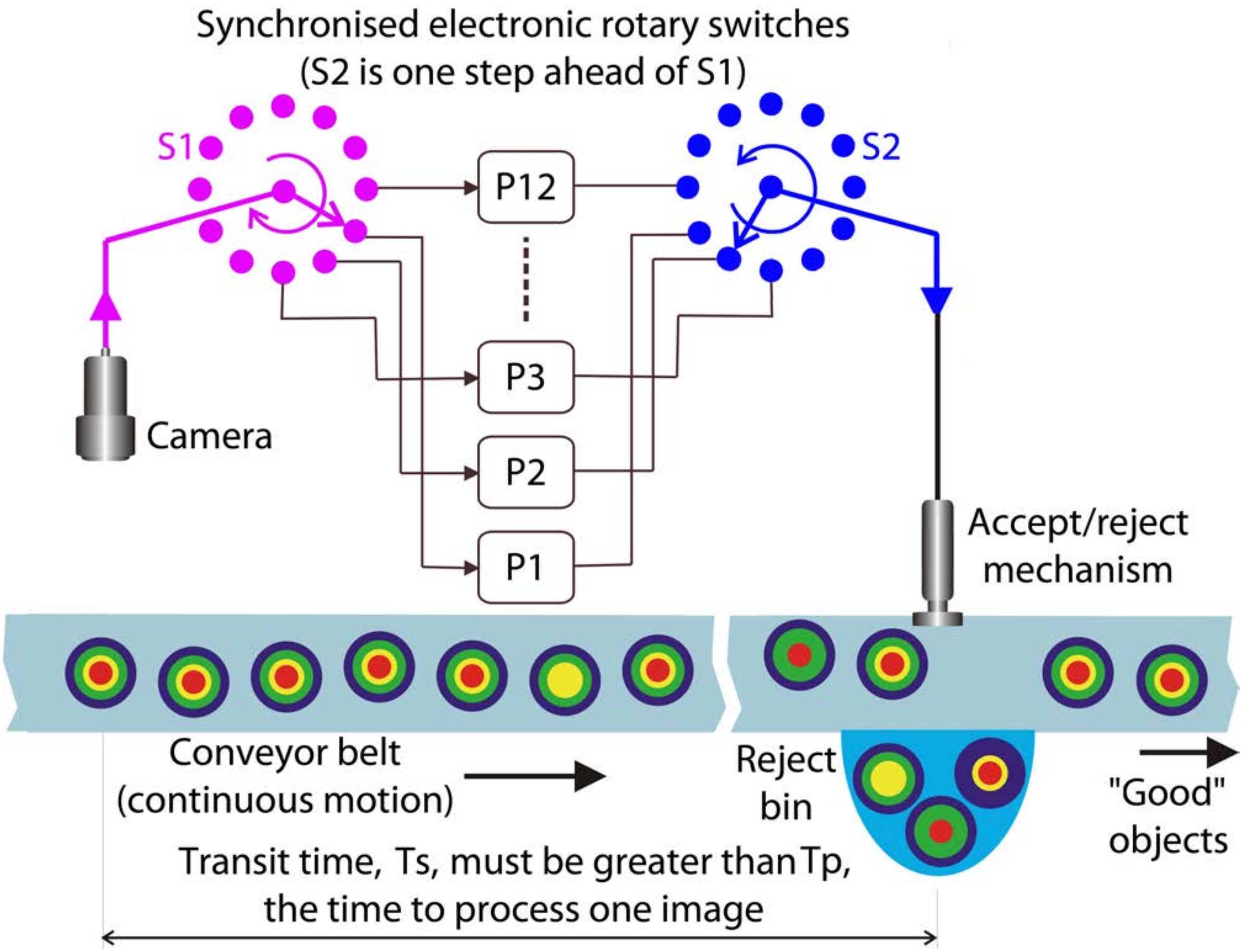


Figure 2.29 A concurrent processor consisting of (12) small "bare bones" processors (P1 - P12). Image data is fed to each of these in turn, using an (electronic) "rotary switch", S1. Each processor runs the same image processing program which takes a time Tp to execute. Switch S2 receives a pass/fail signal from each processor in turn and sends it to the accept/reject mechanism. The rotation of switches S1 and S2 is synchronised and takes a time Ts to perform one cycle. Ts is just a little larger than Tp. If there are N processors (N = 12 here), then at any given time there are N objects lying on the conveyor belt between the camera and the accept/reject point. The belt speed is adjusted so that it takes Ts seconds for an object to travel from the camera to the accept/reject point. The throughput rate is therefore N/Ts objects/second.