1
Basics of Machine Vision

1.1
Digital Images

1.1.1
Grayscale Image

The basic digital image is composed of a two-dimensional array of numbers. Each number in the array represents a value of the smallest visual element, a pixel. The indexed location of the pixel value in the array corresponds to the X and Y locations of the pixel within the image, as measured from the top-left corner. The values of a pixel in an X and a Y location in the digital grayscale image, $f(x,y)$, represent the brightness of the pixel in a range from black to white, as seen in Figure 1.1. Let us assume that total number of pixels are 300 (0–299) and 250 (0–249) in the X and Y locations, respectively. Each image can be represented by the array of size 300×250 that has a value for each pixel.

Each image pixel value is related to the brightness of the image at that specific location. For a given camera device, the maximum value recorded for the image pixels is generally related to a characteristic of the camera referred to as the bit depth. For example, if bit depth is **k**, then there will be as much as 2^k levels of brightness that can be defined. For example, if the bit depth is 8 bits, then a pixel can have 256 values (2^8) in the range between 0 and 255.

A grayscale image pixel most often only has brightness information that can be represented in 8 bit values and as such the image is often referred to as an 8 bit image. If the pixel value is 0, then it is the most dark (black) image pixel, whereas a value of 255 means the brightest image (white) pixel. For a better understanding, Figure 1.1 shows a magnified portion of an image where the location range of X pixels is 85–91 and Y range is 125–130 within a total of 300×250 pixels in the image. In the case of pixel location of 85 along the X direction and 125 along the Y direction, the image pixel value is $f(85, 125) = 197$, which is closer to 255 than 0 and therefore is rendered closer to bright end of the image scale (white). On the other hand, the value of image pixel where $X = 91$ and $Y = 125$ is 14, which is close to 0 and thus relatively dark (black).
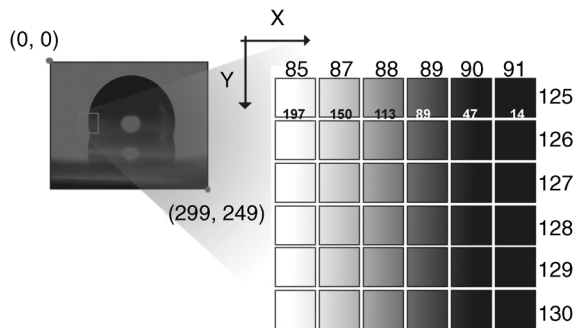
Figure 1.1 Grayscale image.

Due to its simple representation as single pixel values, grayscale images are often used in machine vision applications as a starting point to measure the length or size of an object and to find a similar image pattern via pattern matching. The gray images can be acquired from digital monochrome or color cameras. When the color image is acquired, the color image can easily be converted to a grayscale image by using the color plane extraction function that is provided by NI Vision Development Module.

1.1.2
Binary Image

The most commonly used image format for finding the existence of the object, location, and size information is binary image. The binary image pixel has two digit values, where object has the value of 1 and background has the value of 0 in most cases. Since there are only two values used, it is often called a 1 bit image (bit depth of 1, or 2^1). To make a binary image, the grayscale image is commonly used as a starting point. In general, we use a threshold value to convert a grayscale image to a binary image. In the case that the object of interest in an image is bright against a dark background (the imaged object's pixel value is larger than a chosen threshold value), it is classified as the object (a pixel image value of 1) and if the image value is less than the threshold value, it can be classified as the background (the pixel image value of 0). However, it should be noted that there will be cases where the dark parts of an image may represent the object with the bright part comprising the background.

Once the grayscale image is converted to a binary image, various image processing functions can be used. For example, we can use the particle analysis function from which the size, area, and the center of the object can be easily obtained. Prior to particle analysis, the morphology functions are often used to modify aspects of the image for better or more reliable results. For example, we may want to remove unnecessary parts from the binary image or repair parts

of an object that obviously misrepresents the object in the grayscale to binary conversion. By using the morphology functions in the LabVIEW Vision Development module, we can increase the accuracy of image analysis based on the binary image. Details of this process will be discussed later.

1.1.3
Color Image

Digital color images from digital cameras are usually described by three color values: R (red), G (green), and B (blue). The three color values that represent an image pixel describe the color and brightness of the pixel. In other words, the brightness and color of the pixels in an image obtained from a digital color camera are generally defined by the combination of the R, G, and B values. All possible colors can be represented by these three primary colors. The digital color image is often referred to as a 24 or 32 bit image. Figure 1.2 shows the basic concept of a 32 bit color image. Among four possible 8 bit values in a 32 bit word, we use 8 bits for each of the R, G, and B components. The other 8 bit component is not used. This is due to the computer's natural representation of an integer as a 32 bit number.

Figure 1.3 shows an example of a color image. The total size of the image is 800×600. The X direction has 800 (0–799) pixels and Y direction has 600
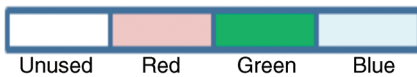

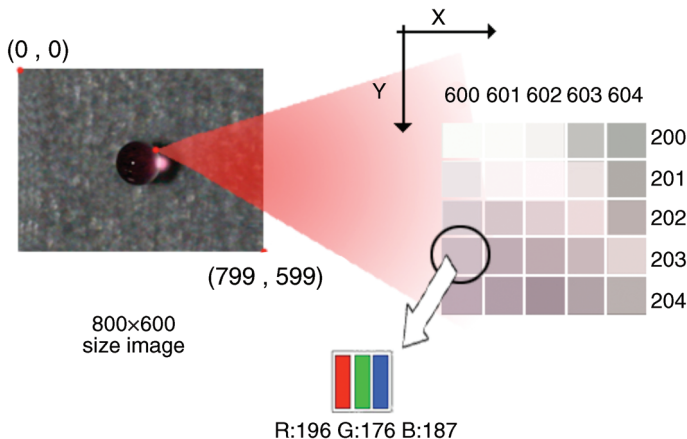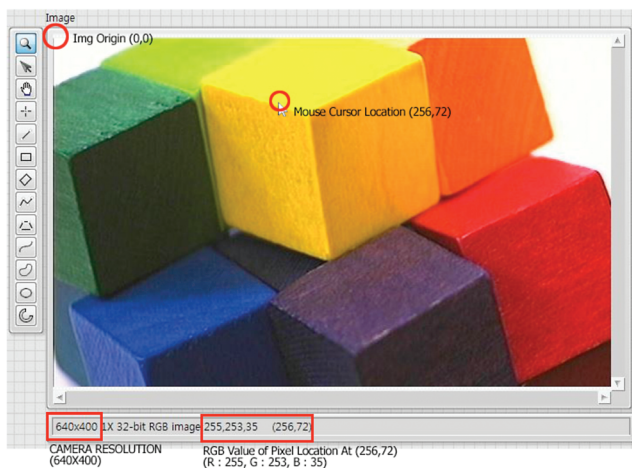**Figure 1.2** 32 bit color image.

Figure 1.3 Color image ($f(x, y) = 0 \leq R \leq 255, 0 \leq G \leq 255, 0 \leq B \leq 255$).

Figure 1.4 Acquired color image.

(0–599) pixels. Each pixel has three component values representing R, G, and B. For example, the image value at X = 600, Y = 203, *f*(600, 203), is R = 196, G = 176, B = 187.

For a better explanation, a USB camera was used to acquire the images via a LabVIEW VI, as shown in Figure 1.4. As seen in the lower part of Figure 1.4, the total size of the image (the number of pixels) is 640 × 480. The pixel location is defined by the X and Y locations, where upper left is (0,0) and lower right is (639,479). Each of the RGB values in a pixel has an 8 bit value, which corresponds to an integer range of 0–255. When we move the mouse cursor over the acquired image, the pixel's RGB values pointed to by the mouse cursor are shown at the bottom of the window. In the example as seen in Figure 1.4, the RGB values at the mouse X/Y image position (257,72) are (255,253,35).

Each pixel color and brightness is the combination of RGB values. For example, R (red) has the range of values between 0 and 255. If the value is close to 0, the R image becomes dark red, which can be seen as black. On the other hand, if the image value of R becomes 255, then the R component becomes the brightest, which is seen as bright red. The green and blue pixel component values have same property. If the R = 255, G = 0, and B = 0, then the pixel appears to be bright red. If all three RGB values are 255, then the pixel appears to be white (bright pixel), whereas if the RGB values are 0, then the pixel becomes dark (black).

One alternative method for color image representation, HSL (hue, saturation, and luminance), can be used instead of RGB (Table 1.1). The three HSL values are also generally represented by 8 bit values for each component. By using proper values of HSL, any color and brightness can be displayed in a pixel.

Table 1.1 The meaning of HSL.

Hue	Saturation	Luminance
Hue defines the color of a pixel such as red, yellow, green, and blue or combination of two of them. It is related to wavelength of a light.	Saturation refers to the amount of white added to the hue and represents the relative purity of a color. If the saturation increases, color becomes pure. If colors are mixed, the saturation decreases. For example, red has higher saturation compared with pink.	Luminance is closely related with the brightness of image. Extracting the luminance values of an HSL color image results in a good conversion of a color image to a grayscale representation.

1.2
Components of Imaging System

Figure 1.5 shows the basic components of imaging systems. Imaging acquisition hardware requires a camera, lens, and lighting source. To get an image from the camera to the computer, we need to select the most appropriate camera communication interface (bus), which connects the camera to the computer. Some cameras require specific types of standardized communication busses integrated into computer interface cards called frame grabbers. Examples of a few standardized frame grabber communication busses are Analog, Camera Link, and Gigabit Ethernet (GigE). Other cameras connect to the computer over more common communication interfaces such as USB, Ethernet, or Fire Wire that are provided as standard configurations in most computers.

Software is also needed to display and extract information from images. In this book, image processing techniques will be described for the purpose of
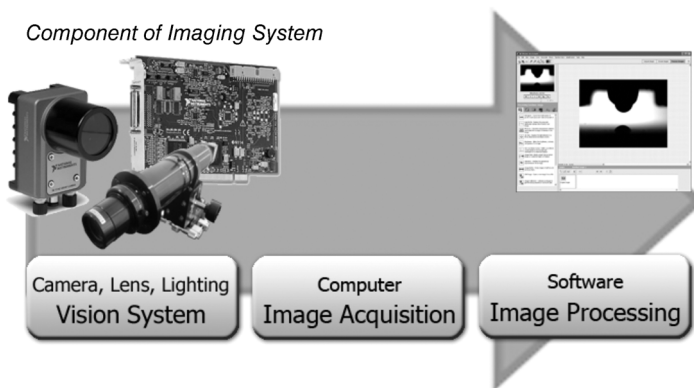
Figure 1.5 Basic component of imaging system.

processing and analyzing the acquired images. While there are a number of software programs that can be used to develop image measurement applications, we will be focusing on methods using the LabVIEW Vision Development module from National Instruments.

1.2.1
Camera

To acquire images, the camera selected must match the requirements suitable to the specifics of imaging task. For this purpose, a brief overview of cameras will be discussed in this section. For better camera selection, we recommend that you should consult with your camera vendor.

Color and Monochrome Camera

If the imaging task can benefit from the additional information provided by the color image of an object or set of objects, a color camera is required. However, one needs to take into account the increased data set size (possibly 4x) and complexity required for processing color images. Therefore, a decision needs to be made as to whether a color camera is required based on the application. As an example, it may be better to enhance the appearance of an object of a specific color by using a monochrome camera in combination with a color filter that may increase the contrast of the object in a grayscale image.
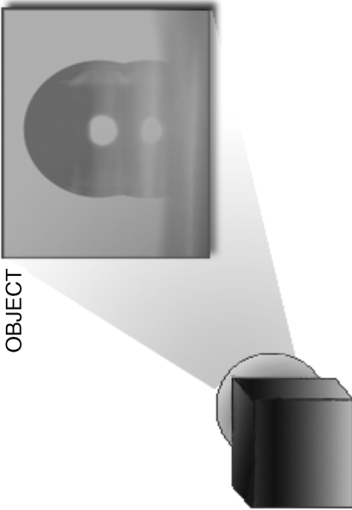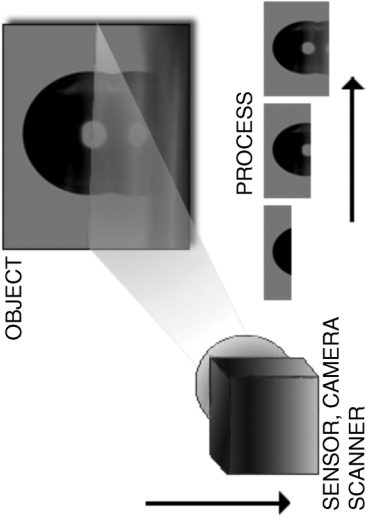
Frame Rate

Frame rate means the number of images (or "frames") acquired per second. The unit is frame per second (fps). The frame rate of most cameras for vision measurement purpose is about 30 fps. This is an historical value based on the development of the television in the United States where the frame rate was determined by half the alternating power current frequency of 60 Hz. On the other hand, when there are needs for high-speed real-time monitoring, the proper selection of high frame rate camera hardware may be required.

Area Scan Camera and Line Scan Camera

Digital cameras can be classified as area scan cameras or line scan cameras according to the scan method. Line scan cameras use one-dimensional sensor arrays that acquire a one-dimensional image in a single frame. Area scan cameras have an image sensor that can acquire a two-dimensional image in a single frame. In most general vision applications, area scan cameras are used. However, in case of inspecting moving object or where the camera is moving, the line scan camera may be best to use for fast inspection. The principle of line scan camera is quite similar to the document scanner. If the object is moving in a perpendicular direction relative to the sensor array in a line scan camera, it can acquire one- or two-dimensional images, as seen in the line camera in Table 1.2.

Table 1.2 Comparison between area scan and line scan camera.

Area scan camera	Line scan camera
AREA SCAN	LINE SCAN

OBJECT

SENSOR, CAMERA SCANNER

PROCESS

OBJECT

SENSOR, CAMERA SCANNER

| Two-dimensional array sensors | One-dimensional array sensors |
| Commonly used | Image acquisition by moving camera or object |

Frame Trigger

The trigger signals from (or to) the camera or frame grabber can be used to synchronize image acquisition with respect to external measurement device, lighting application, or motion of a stage.

Image Resolution

The image resolution is important because it is related to the accuracy of the vision measurement. Resolution is related to the lens image magnification and camera resolution (pixel size and number of pixels). The zoom lens can be used to increase the image resolution. However, increasing the zoom factor of a lens often results in a narrowing of the field of view (FOV), which is defined as the physical dimensions that the image represents. It is often recommended to use a high-resolution camera with more pixels in the image sensors. However, the choice of a high-resolution camera often increases the costs of an imaging system and impacts the computational requirement due to the increased image data set size. So, the proper camera needs to be selected according to imaging task requirements. As a general rule of thumb, two or more pixels are required to detect any defects and more than 10 pixels are required to measure the size of an object. However, it should be noted that the resolution requirement differs according to the specific requirements of the application.

Historically, inexpensive area scan cameras with 30 fps have 640×480 pixel sensor arrays. However, there are cameras available with different numbers of pixel sensor arrays. Table 1.3 shows the example of some commercially available cameras.

In general, the number of camera pixel sensors and the field of view of the camera/lens system are critical factors for determining the image resolution. It should be noted that FOV should be large enough to measure the object of interest. If you know the FOV, you can use the following equation to determine the image resolution:

$$\text{Resolution} = \left(\frac{\text{FOV}}{\text{number of camera pixels in one direction}} \right) \times 2$$

Table 1.3 Example of camera sensors in a camera.

Area scan camera	Line scan camera
640×480	512×1
752×582	1024×1
1024×768	2048×1
1024×1024	4068×1
1280×960	6144×1
1360×1024	8192×1
1620×1220	$12\,228 \times 1$
1920×1080	
2048×2048	
4872×3248	

If the FOV of horizontal direction is 64 mm and the number of sensors in the X direction is 640, the image resolution can be calculated:

$$\text{Image resolution} = \left(\frac{64}{640}\right) \times 2 = 0.2 \, \text{mm}$$

Camera Sensor Size

Figure 1.6 shows the relationship between camera sensor size and FOV. The sensor size differs according to the number of pixels and the size of a pixel. As seen in Figure 1.6, the sensor's pixel size is important because it is directly related to the selection of the lens.

Area of Interest

Area of interest (AOI) is used when the fast image acquisition is required by acquiring a part of an image from the camera sensors (Figure 1.7). A similar concept is region of Interest (ROI). ROI differs from AOI in that the former uses the software algorithm to process the part of acquired images, whereas the latter is more hardware-based concept for image acquisition. The concept and application of ROI will be discussed later.
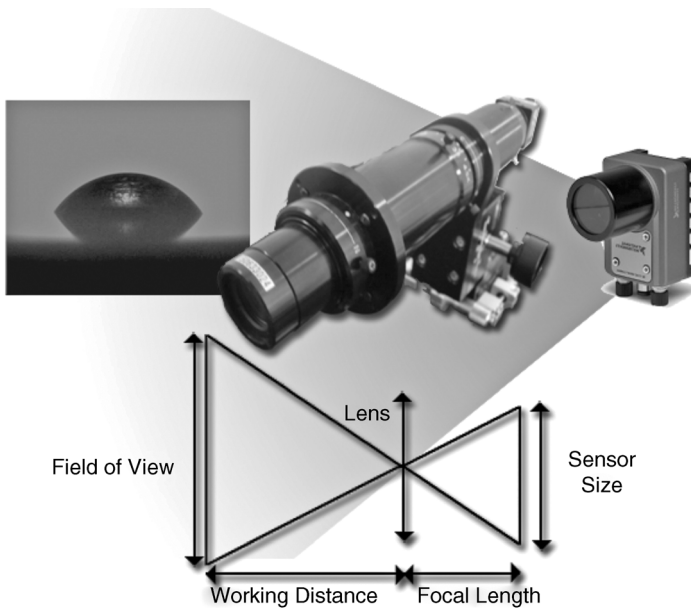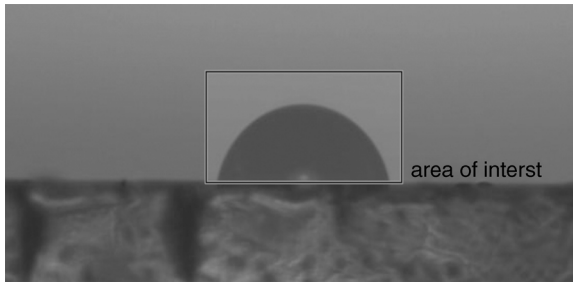

Figure 1.6 Camera lens selection.

Figure 1.7 Area of interest.

1.2.2
Camera Bus: The Method to Connect PC and Camera

To acquire an image from a camera, several vision acquisition interface methods have been developed, including analog, camera link, USB, IEEE 1394, and GigE. To determine the proper type of camera bus, we need to compare the camera bus' capabilities according to the specific application of interest. Once we decide the right camera bus, the National Instruments (NI) Web site (http://www.ni.com/camera) can be referred to select an appropriate camera.

You may call NI technical center to get information on the proper image acquisition board from NI products according to your application.

Analog Camera
To acquire images from the analog signal produced by an analog camera, BNC or RCA cables are commonly used, as seen in Figure 1.8.

There is no power source provided for the camera in the analog camera bus. Therefore, analog cameras generally require an external power source (such as

Figure 1.8 Analog camera connected to an analog frame grabber card with a BNC cable. Video standards for color and monochrome analog camera are summarized in Table 1.4.

Table 1.4 Standard analog video.

	Standard	Number of image sensors	Frame rate (fps)	Comments
Color	NTSC (National Television Systems Committee)	640×480	29.97	North America, Japan
	PAL (Phase Alternative Line)	768×576	25	Europe
Monochrome	RS170 (Electronic Industries Association)	640×480	30	North America
	CCIR (Consultative Committee for International Radio)	768×576	25	Europe

12V DC). Analog cameras also require an analog frame grabber to convert the composite video analog signal to digital image.

Camera Link

One standard for a high-speed video bus is known as Camera Link, which was determined by the Automated Imaging Association (AIA). The standard defines the cable between the camera and a frame grabber, the connectors, and the signals and their functions (Figure 1.9).

Figure 1.9 Camera link cable.

Cameras designed to the Camera Link standard work with all the Camera Link-specific frame grabbers. Camera Link is a specially designed high-speed digital bus. Some base-priced Camera Link cameras can acquire 1 megapixel image at 50 fps. Medium- and high-performance cameras acquire 510 and 680 MB/s, respectively. Some higher performance cameras can acquire 1280×1024 images at 500 fps.

The Camera Link camera bus is designed for middle- or high-end applications and the price of these imaging systems reflects this capability. In addition, Camera Link cameras require a frame grabber that is capable of high-speed processing. The price of a Camera Link frame grabber is more than that of analog frame grabbers. However, as the digital camera systems become of more standard and the analog systems less standard, it would be harder to find electronics to support analog systems.

National Instruments requires Camera Link cameras to have special camera description files, which have information on image acquisition and the communication method that can be used by the NI software to acquire the camera image. Camera description files can be found from http://www.ni.com/cameras.

USB Camera

The initial USB 1.1 standard did not have enough speed or bandwidth to support the data requirements of most imaging applications. However, the USB 2.0 standard has increased bandwidth capable of video streaming and comparable speeds to IEEE 1394a. USB 3.0 has even greater capability.

The advantage of USB cameras is that they are relatively inexpensive and do not require a frame grabber. As a result, USB cameras are convenient for research purposes and even in industry when there are cost issues and special functions may not be required (e.g., triggering, etc.).

IEEE 1394

Historically, the initial image acquisition speed of FireWire, or IEEE 1394a (Figure 1.10), was much faster than USB 1.1. So, due to the high bandwidth requirements of cameras, it has been a standard in many vision acquisition systems. Since the 1394 camera does not require the frame grabber and power can be provided from the cable, the vision system can be simplified.

The drawback to 1394 imaging systems is the relatively higher price compared with USB camera systems. In addition, due to the development of USB 2.0 and

Figure 1.10   IEEE 1394.

USB 3.0, the communication speed of USB cameras is now becoming comparable with that of the 1394 cameras.

Nonetheless, there are several merits in 1394 camera compared with USB. The IEEE1394 camera can work independently and can communicate with other devices without a computer. In comparison, USB cameras need a master controller and are required to operate under the master control, which is usually supplied in the form of a computer. Also, it is known that the 1394 camera systems are generally considered more reliable in an industrial environment compared with USB camera.

Gigabit Ethernet
GigE cameras (Figure 1.11) use gigabit Ethernet (LAN cable) for real-time data and image transfer to computer. There is no need for an additional frame grabber with the GigE camera. As a result, high-speed and low-cost image acquisition is possible by using GigE camera. The GigE camera can use very long camera cables up to 100 m. However, external power is still required.
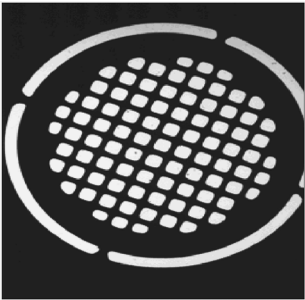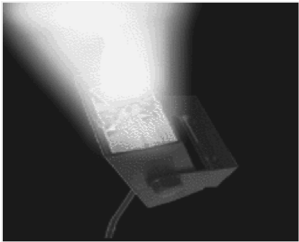
1.2.3
Lens

The selection of an appropriate lens is crucially important for any application being considered. The choice of lens has significant effects on the FOV, working

Figure 1.11   GigE camera.

Table 1.5 Various lightings.

Ring light	Back light	Strobe light	Diffused lighting
Ring-shaped LED arrays or fiber optic ring lights can be used to focus the light on the object located in the left of the ring for clear images	By using the back light, the object shape and size can be investigated. Relatively small amount of light is required. It is a useful method for inspecting an object's outer shape, but there can be objectionable effects such as light diffraction around the object	This method can be used to obtain images of object frozen in motion

The light needs to be synchronized with respect to the image capture timing and position of the object of interest in motion | The light passes through a diffused plate such that it can result in uniform lighting on the area of interest |
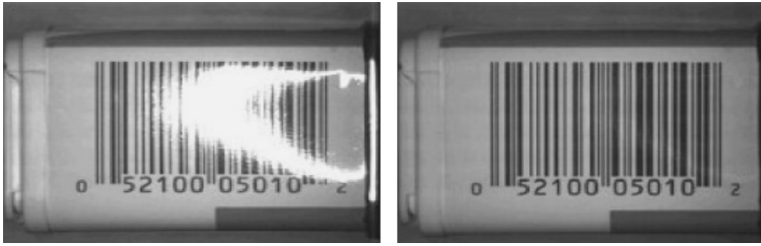
Figure 1.12 The importance of lighting.

distance, and optical image resolution at the camera's sensor. To select a proper lens, the focal length of the lens is often used. Focal length is defined by distance between lens and the image plane at the sensor in the camera. Figure 1.6 shows the relationship among focal length, FOV, sensor size, and working distance. Here, the working distance means the distance between lens and object to measure. If you know the FOV, sensor size, and working distance, you can calculate focal length of lens by the following equation:

$$\text{Focal length} = \frac{\text{sensor size} \times \text{working distance}}{\text{FOV}}$$

1.2.4
Lighting

The main purpose of lighting is to differentiate the background from the object to be measured by providing contrast. The contrast means the light intensity difference between the background and the object to measure. To extract image information for vision analysis, the imaged object should have enough light intensity difference to distinguish it from its imaged surroundings. To optimize the contrast in the acquired images, proper lighting is essential prior to image acquisition. Figure 1.12 shows an example of the importance of corrected lighting. If the lighting is inadequate, we cannot get the required information from the acquired image.

As a power source for lighting, DC or high-frequency lighting is commonly used. In specific applications, a strobe light synchronized with motion of objects of interest can be used very effectively. There are many lighting tricks that can be applied in specific applications. Table 1.5 shows the few examples.