FPT Education

FPT UNIVERSITY

PART I
IMAGE PROCESSING

INSTRUCTOR: DR. NGUYEN NGOC TRUONG MINH

SCHOOL OF ELECTRICAL ENGINEERING, INTERNATIONAL UNIVERSITY (VNU-HCMC)

Ho Chi Minh City, June 2023

LECTURE IV – IMAGE SENSING AND ACQUISITION

INSTRUCTOR: DR. NGUYEN NGOC TRUONG MINH

SCHOOL OF ELECTRICAL ENGINEERING, INTERNATIONAL UNIVERSITY (VNU-HCMC)

Ho Chi Minh City, June 2023

LECTURE CONTENT

- What are the main parameters involved in the design of an image acquisition solution?
- How do contemporary image sensors work?
- What is image digitization and what are the main parameters that impact the digitization of an image or video clip?
- What is sampling?
- What is quantization?
- How can I use MATLAB to resample or requantize an image?
- Chapter Summary – What have we learned?
- Problems

4.1 INTRODUCTION

- *Remind: an image as a two-dimensional (2D) representation of a real-world, three-dimensional (3D) object or scene and indicated that the existence of a light source illuminating the scene is a requirement for such an image to be produced. The concept of a digital image as a representation of a two-dimensional image using a finite number of pixels was also introduced.*
- In this lecture, we will look at relevant issues involved in *acquiring and digitizing an image*, such as *the principles of image formation* as a result of reflection of light on an object or scene, the sensors typically used to capture *the reflected energy*, and the technical aspects involved in selecting *the appropriate number of* (horizontal and vertical) *samples and quantization levels for the resulting image*.
- In other words, we will present *the information necessary* to understand how we go from real-world scenes to 2D digital representations of those scenes.

4.1 INTRODUCTION

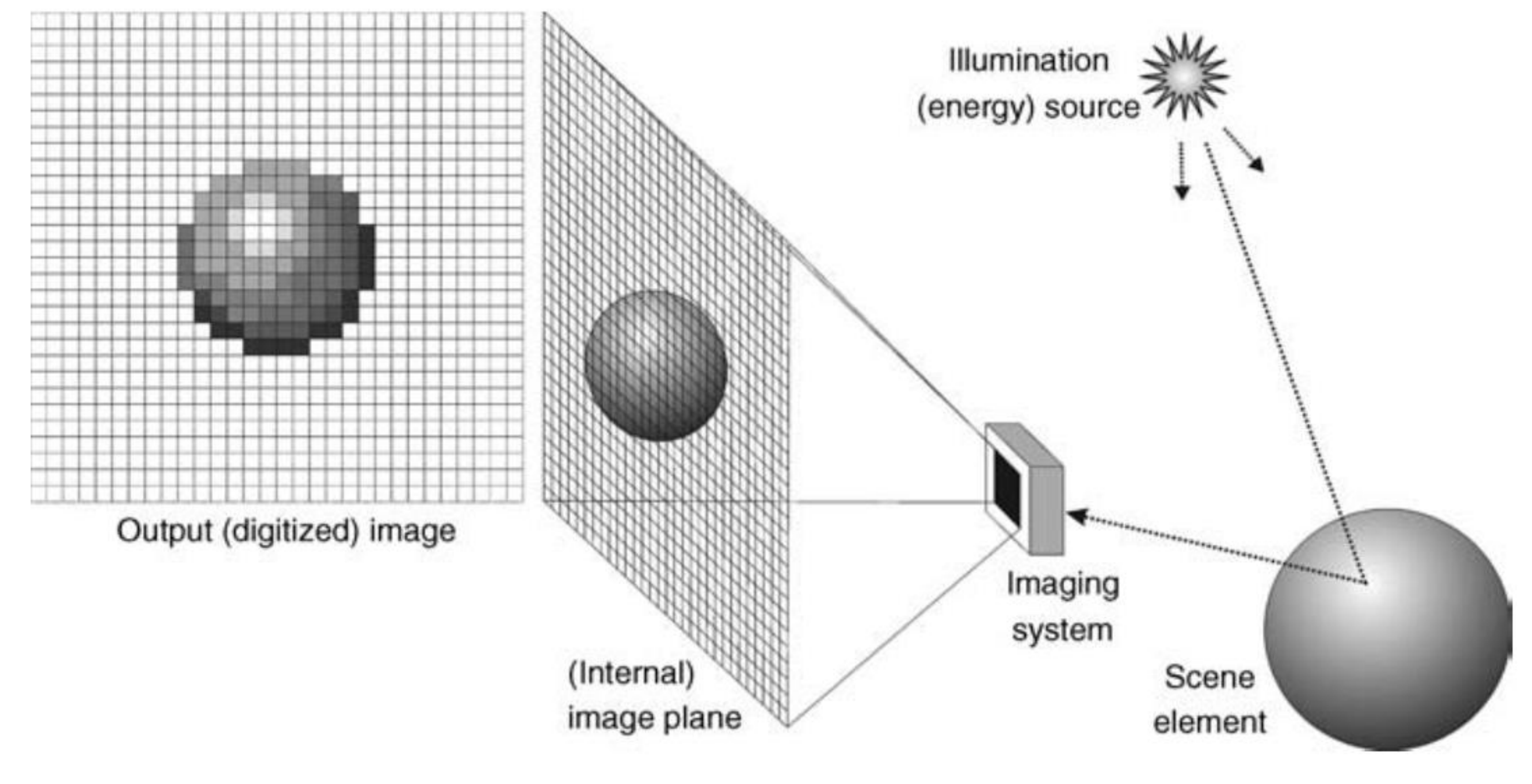


FIGURE 5.1 Image acquisition, formation, and digitization. Adapted and redrawn from [GW08].

4.2 LIGHT, COLOR, AND ELECTROMAGNETIC SPECTRUM

- The existence of *light*—or other forms of electromagnetic (EM) radiation—is *an essential requirement for an image to be created, captured, and perceived*.

- In this section, we will look at basic concepts related to light, the perception of color, and the electromagnetic spectrum.

4.2.1 LIGHT AND ELECTROMAGNETIC SPECTRUM

- Light can be described in terms of electromagnetic waves or particles, called *photons*. A photon is a tiny packet of vibrating electromagnetic energy that can be characterized by its wavelength or frequency.
- Wavelength (λ) and frequency (f) are related to each other by the following expression:

$$\lambda = \frac{v}{f}$$

where v is the velocity at which the wave travels, usually approximated to be equal to the speed of light (c): 2.998x10^8 m/s.

4.2.1 LIGHT AND ELECTROMAGNETIC SPECTRUM

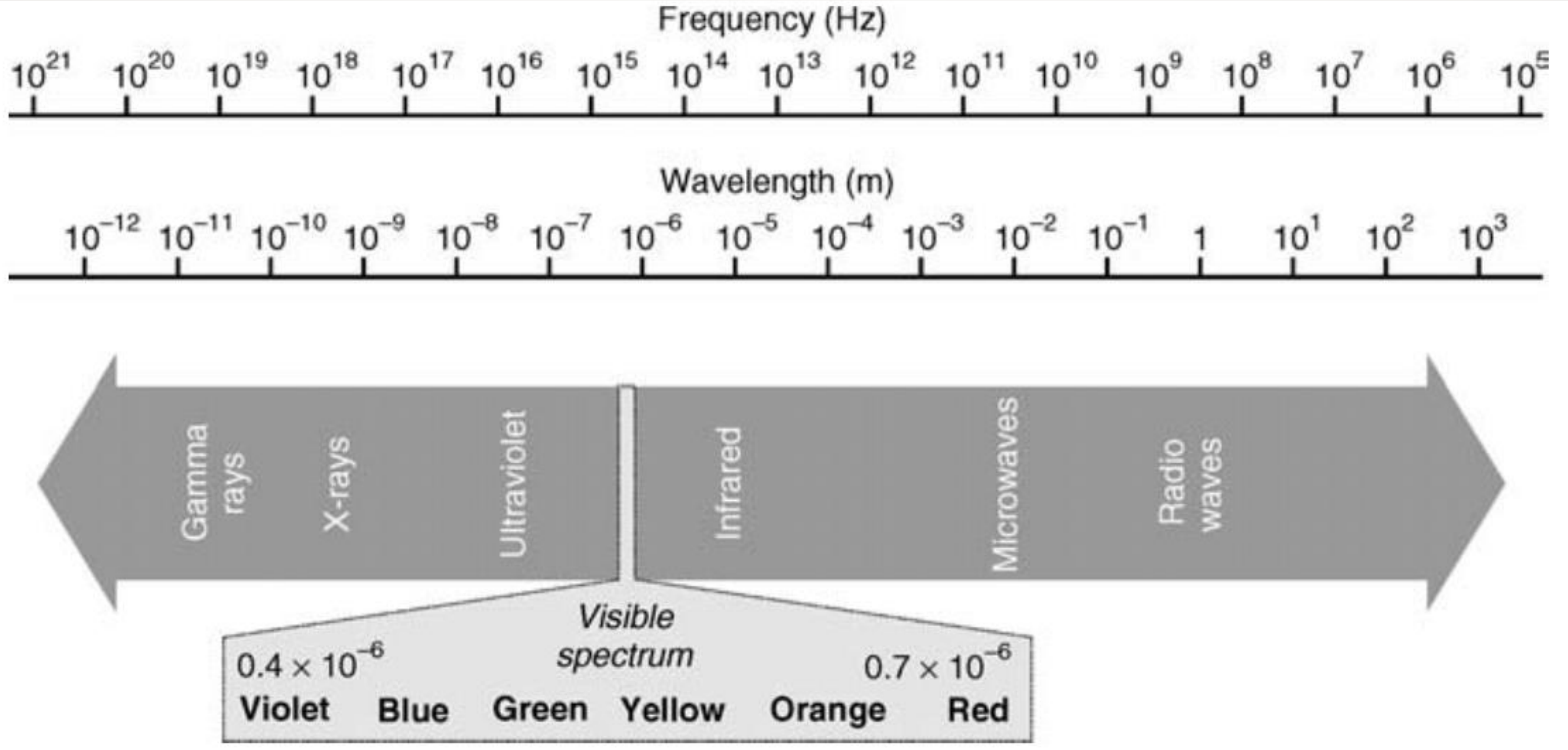
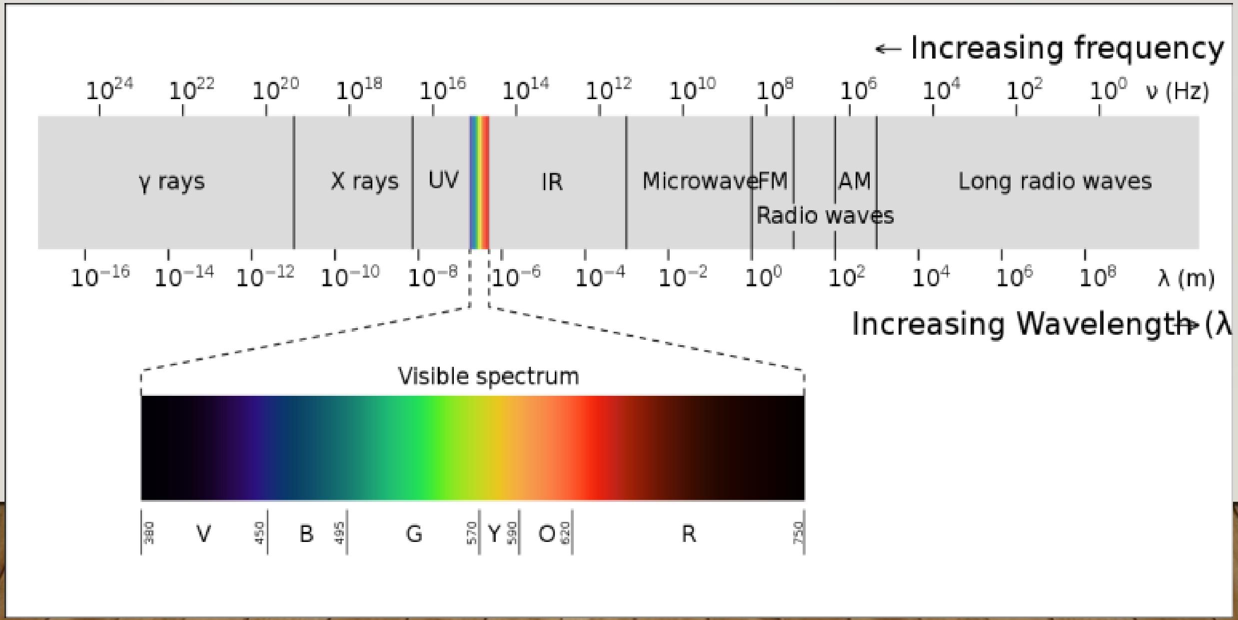


FIGURE 5.2 Electromagnetic spectrum.

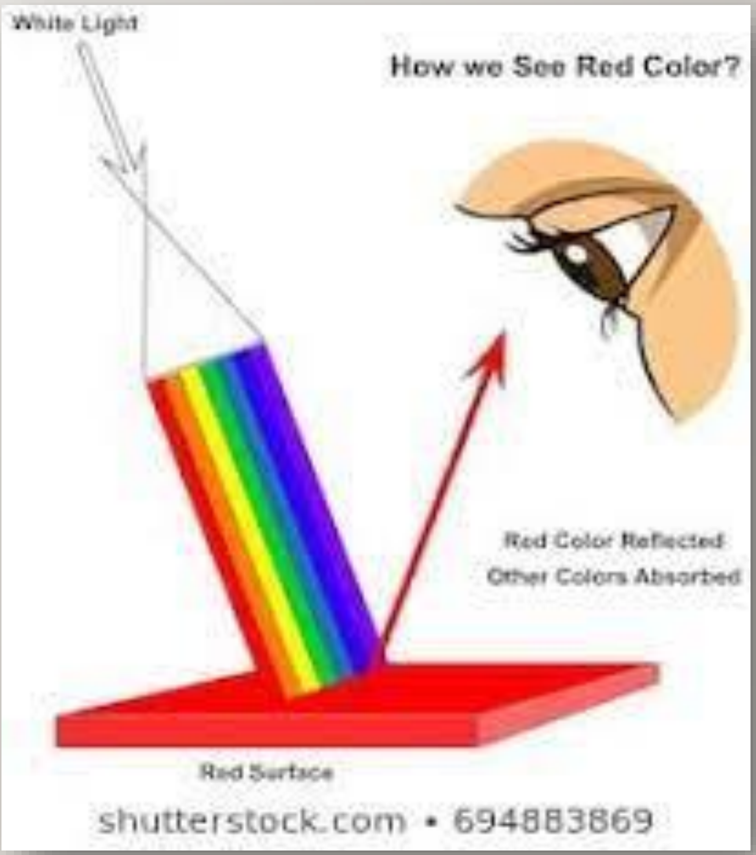
4.2.1 LIGHT AND ELECTROMAGNETIC SPECTRUM

- The human visual system (HVS) is sensitive to photons of wavelengths *between 400 and 700 nm*. As shown in Figure 5.2, this is a fairly narrow slice within the EM spectrum, which ranges from *radio waves* (*wavelengths of 1m or longer*) at one end to *gamma rays* (*wavelengths of 0.01 nm or shorter*) at the other end.



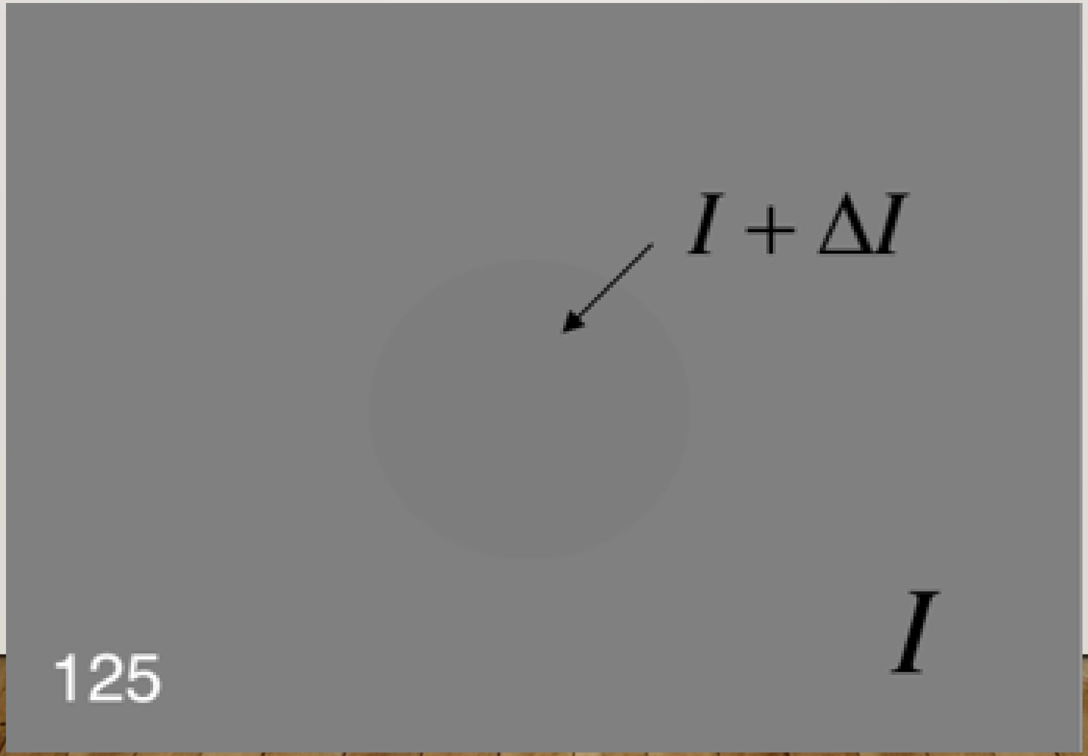
4.2.1 LIGHT AND ELECTROMAGNETIC SPECTRUM

- Even though much of the progress in image processing has been fostered by work on *images outside the visible spectrum*, captured with specialized sensors, we will focus exclusively on *images within the visible range of the EM spectrum*.
- Light is the preferred energy source for most imaging tasks because it is *safe, cheap, easy to control and process with optical hardware*, *easy to detect using relatively inexpensive sensors*, and *readily processed by signal processing hardware*.



4.2.2 BRIGHTNESS DISCRIMINATION EXPERIMENT

- Visibility threshold $\Delta I/I \cong 1..2\%$ ("Weber fraction" / "Weber Law")
- I is luminance, measured in cd/m²



4.2.2 BRIGHTNESS DISCRIMINATION EXPERIMENT

4.2.3 TYPES OF IMAGES

- Images can be classified into three categories according to the type of interaction between *the source of radiation*, *the properties of the objects involved*, and *the relative positioning of the image sensor*.

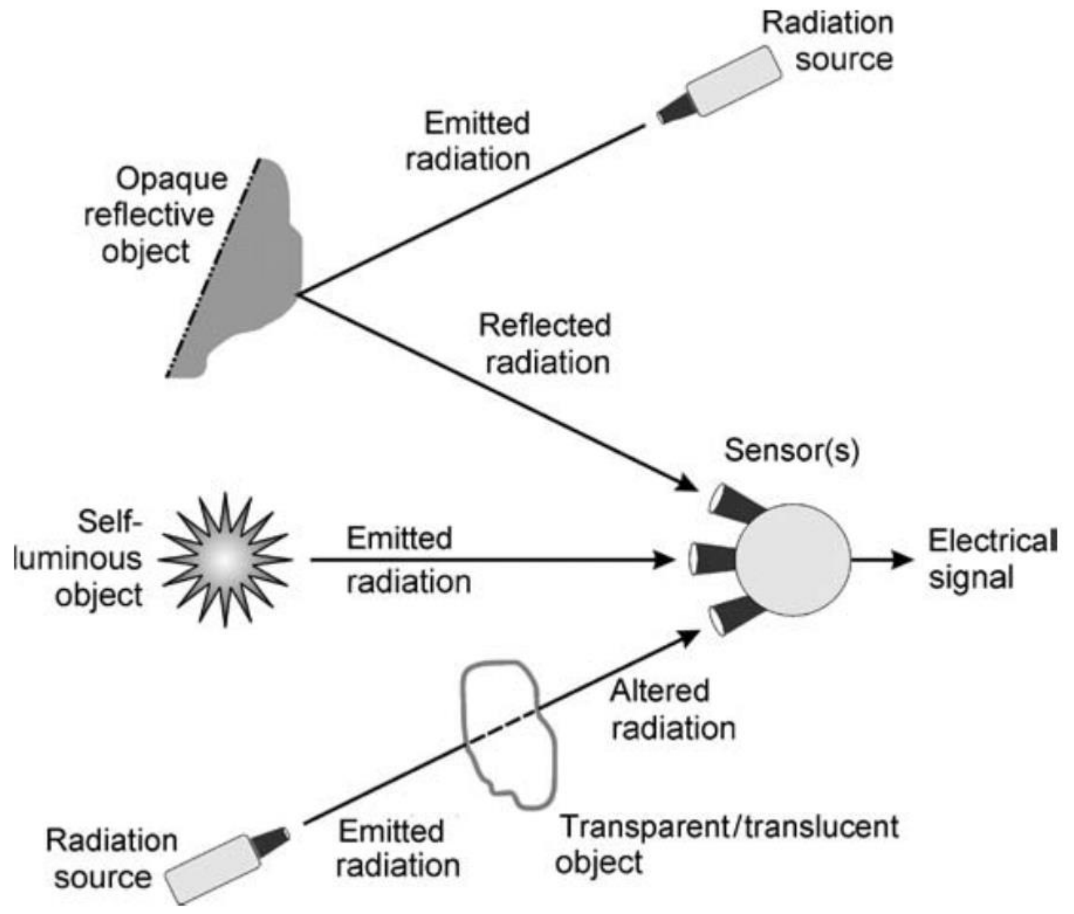


FIGURE 5.3 Recording the various types of interaction of radiation with objects and surfaces. Redrawn from [Bov00a].

4.2.3 TYPES OF IMAGES

- *Reflection Images*: These are the result of radiation that has been reflected from the surfaces of objects. The radiation may be ambient or artificial. Most of the images we perceive in our daily experiences are reflection images. The type of information that can be extracted from reflection images is primarily about surfaces of objects, for example, their shapes, colors, and textures.
- *Emission Images*: These are the result of objects that are self-luminous, such as stars and light bulbs (both within the visible light range), and—beyond visible light range—thermal and infrared images.
- *Absorption Images*: These are the result of radiation that passes through an object and results in an image that provides information about the object's internal structure. The most common example is X-ray image.

4.2.4 LIGHT AND COLOR PERCEPTION

- Light is *a particular type of EM radiation* that can be sensed by the human eye.
- Colors perceived by humans are determined by *the nature of the light reflected by the object*, which is a function of the spectral properties of the light source as well as the absorption and reflectance properties of the object.
- Sir Isaac Newton "*A beam of sunlight passing through a prism undergoes decomposition into a continuous spectrum of components. Each of these components produces a different color experience, ranging from what we call red at one end to violet at the other*" ⇨ Newton's experiments and theories gave fundamental insights into the physical properties of light.
- More important, the "colors" were not in the light itself but in the effect of the light on the visual system.

4.2.4 LIGHT AND COLOR PERCEPTION

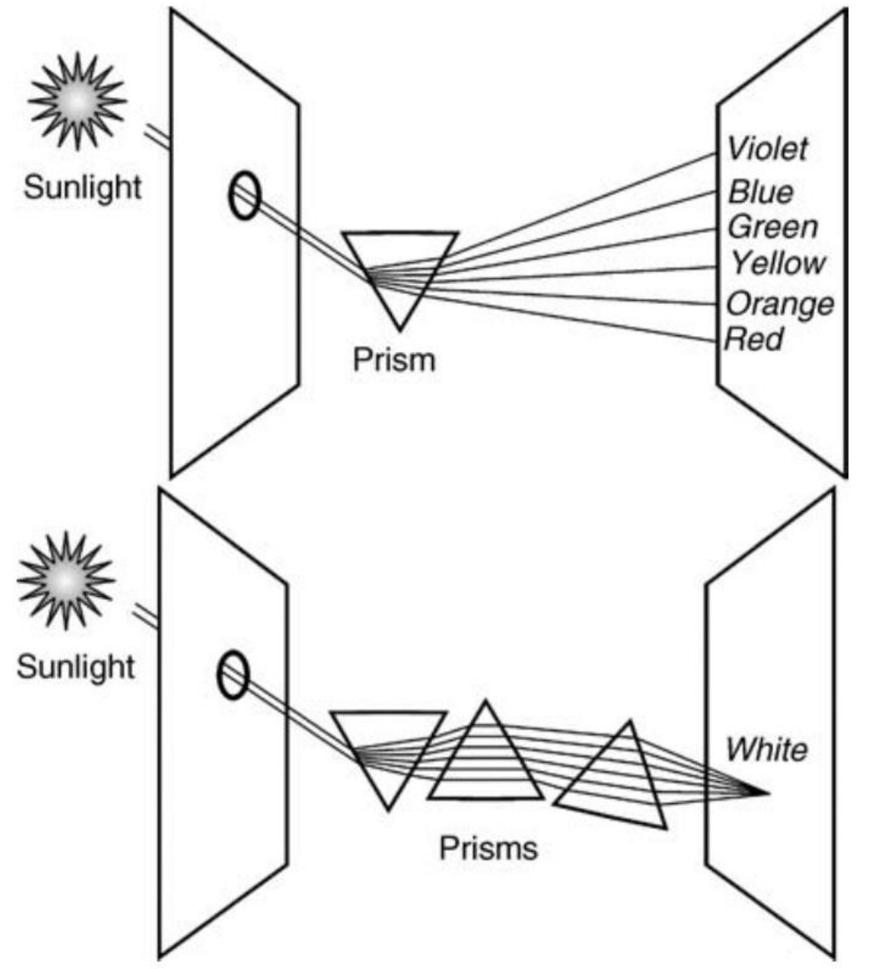


FIGURE 5.4 Newton's prism: many "colors" in the sunlight.

4.2.4 LIGHT AND COLOR PERCEPTION

- The radiance (physical power) of a light source is expressed in terms of its *spectral power distribution* (SPD).
- Figure 5.5 shows examples of SPDs of physical light sources commonly found in imaging systems: sunlight, tungsten lamp, light-emitting diode (LED), mercury arc lamp, and helium–neon laser.
- The human perception of each of these light sources will vary—from the yellowish nature of light produced by tungsten light bulbs to the extremely bright and pure red laser beam.

4.2.4 LIGHT AND COLOR PERCEPTION

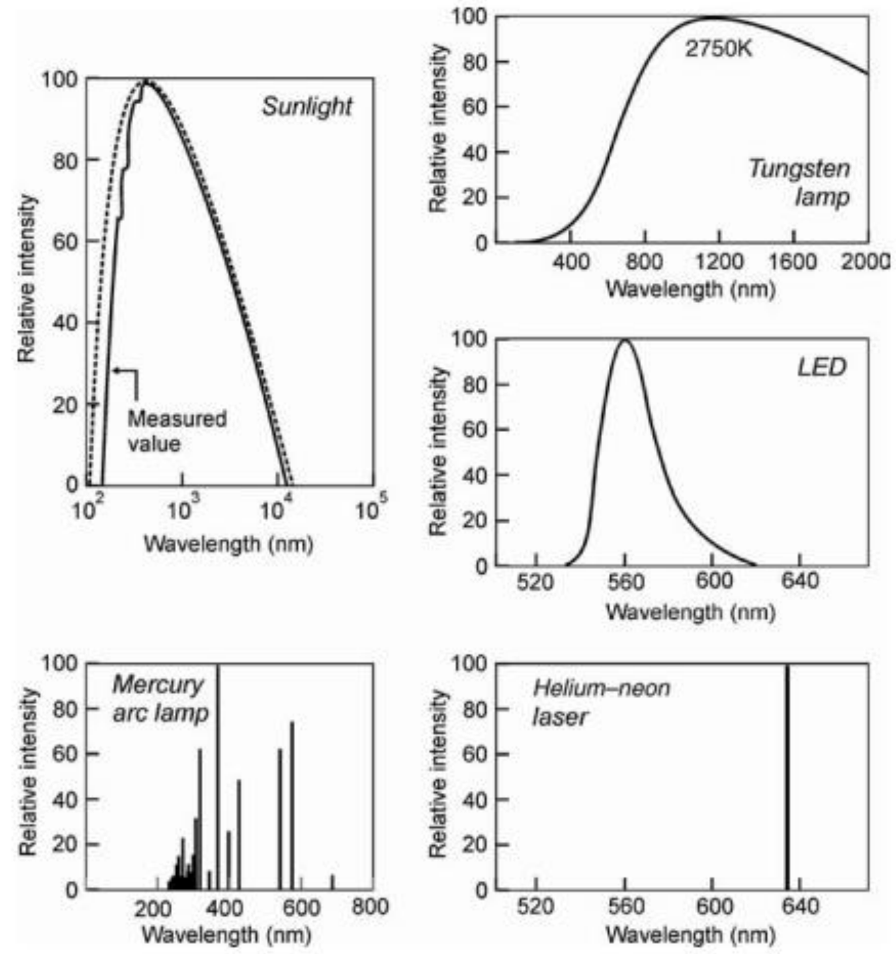


FIGURE 5.5 Spectral power distributions of common physical light sources. Redrawn from [Pra07].

4.2.5 COLOR ENCODING AND REPRESENTATION

- Color can be encoded using *three numerical components* and appropriate spectral weighting functions.
- *Colorimetry* is the science that deals with the quantitative study of color perception. It is concerned with the representation of tristimulus values, from which the perception of color is derived.
- The simplest way to encode color in cameras and displays is by using the red (R), green (G), and blue (B) values of each pixel.

4.2.5 COLOR ENCODING AND REPRESENTATION

- Human perception of light—and, consequently, color—is commonly described in terms of three parameters:
  - **Brightness**: *The subjective perception of (achromatic) luminous intensity, or "the attribute of a visual sensation according to which an area appears to emit more or less light".*

4.2.5 COLOR ENCODING AND REPRESENTATION

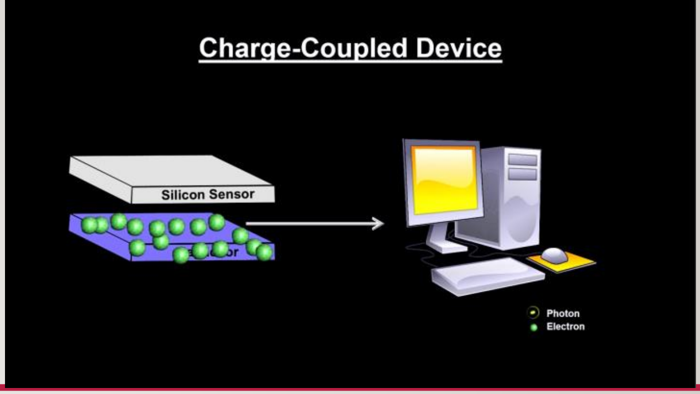
- Human perception of light—and, consequently, color—is commonly described in terms of three parameters:
  - ***Hue***: *"The attribute of a visual sensation according to which an area appears to be similar to one of the perceived colors, red, yellow, green and blue, or a combination of two of them". From a spectral viewpoint, hue can be associated with the dominant wavelength of an SPD.*

4.2.5 COLOR ENCODING AND REPRESENTATION

- Human perception of light—and, consequently, color—is commonly described in terms of three parameters:
  - *Saturation: "The colorfulness of an area judged in proportion to its brightness", which usually translates into a description of the whiteness of the light source. From a spectral viewpoint, the more an SPD is concentrated at one wavelength, the more saturated will be the associated color. The addition of white light, that is, light that contains power at all wavelengths, causes color desaturation.*

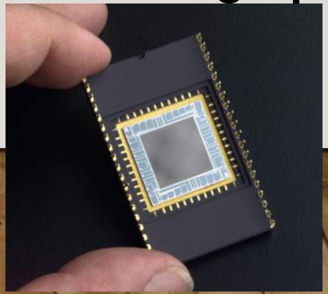
4.3 IMAGE ACQUISITION

- In this section, we describe the basics of image acquisition and its two main building blocks: *the image sensor* and *the optics associated with it*.



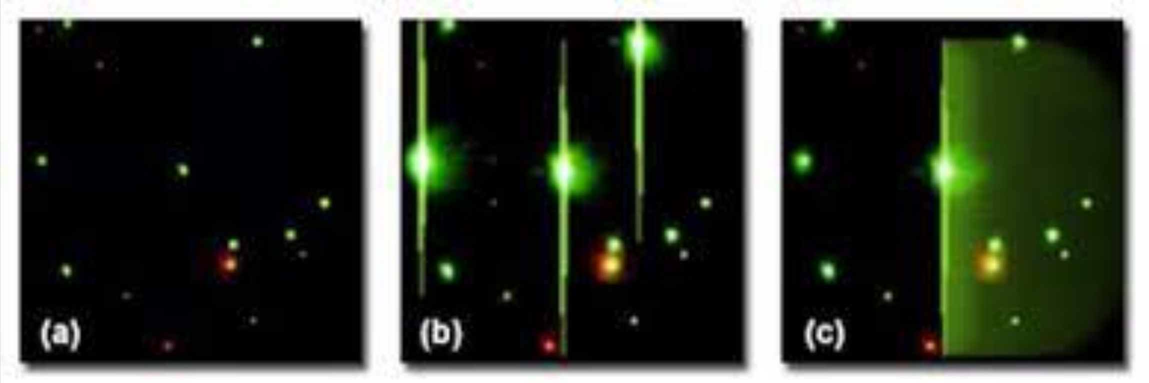
4.3.1 IMAGE SENSORS

- The main goal of an image sensor is *to convert EM energy into electrical signals* that can be processed, displayed, and interpreted as images.
- The way this is done varies significantly from one technology to another.
- Two of the most popular and relatively inexpensive devices used for image acquisition are *the digital camera* and *the flatbed scanner*. Cameras typically use 2D (area) CCD (charge-coupled devices) sensors, whereas scanners employ 1D (line) CCDs that move across the image as each row is scanned.
- A CCD sensor is made up of an array of light-sensitive cells called *photosites*, manufactured in silicon, each of which produces a voltage proportional to the intensity of light falling on them.



4.3.1 IMAGE SENSORS

- A saturated photosite can overflow, corrupting its neighbors and causing a defect known as *blooming*.
- The *nominal resolution* of a CCD sensor is the size of the scene element that images to a single pixel on the image plane. For example, if a 20cm x 20cm square sheet of paper is imaged to form a 500 x 500 digital image, then the nominal resolution of the sensor is 0.04cm.



Images without blooming (a) and with vertical (b) and horizontal (c) blooming

4.3.1 IMAGE SENSORS

- The *field of view* (FOV) of an imaging sensor is *a measure of how much of a scene it can see*, for example, 10cm x 10cm. Since this may vary with depth, it is often more meaningful to refer to the angular field of view, for example, 55º x 40º.
- A CCD camera sometimes plugs into a computer board, called *frame buffer*, which contains fast access memory (typically 0.1ms per image) for the images captured by the camera.

4.3.1 IMAGE SENSORS

- In single-CCD cameras, colors are obtained by using a *tricolor imager* with different photosensors for each primary color of light (*red*, *green*, and *blue*), usually arranged in a Bayer pattern (Figure 5.6). In those cases, each pixel actually records only one of the three primary colors; to obtain a full-color image, a *demosaicing* algorithm—which can run inside the actual camera, before recording the image in JPEG format, or in a separate computer, working on the raw output from the camera—is used to interpolate a set of complete R, G, and B values for each pixel.
- More expensive cameras use three CCDs, one for each color, and an *optical beam splitter*.

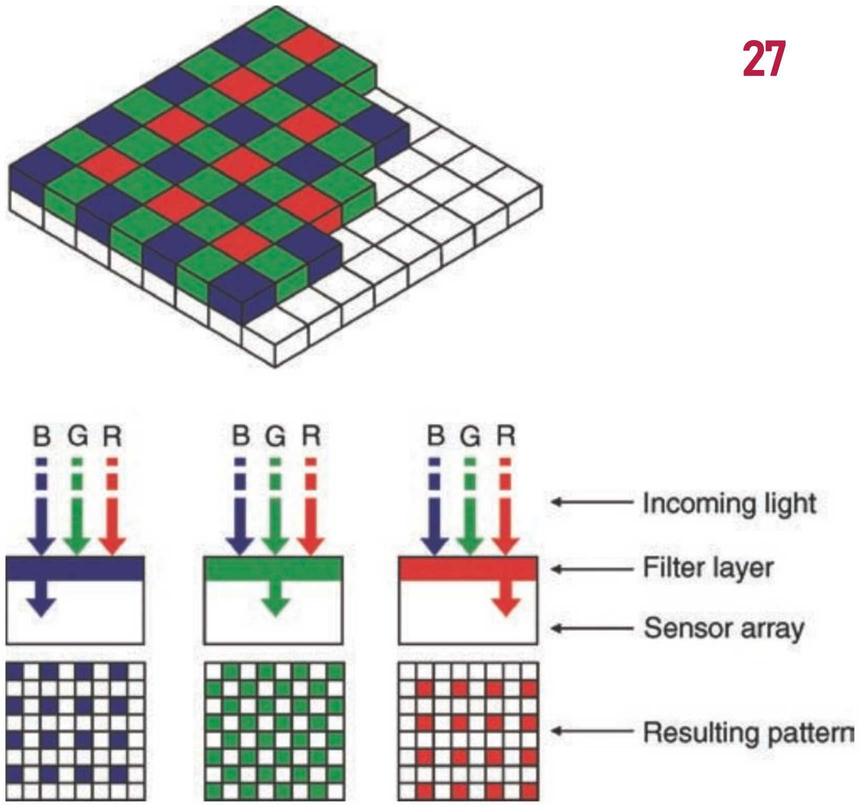


FIGURE 5.6 The Bayer pattern for single-CCD cameras.

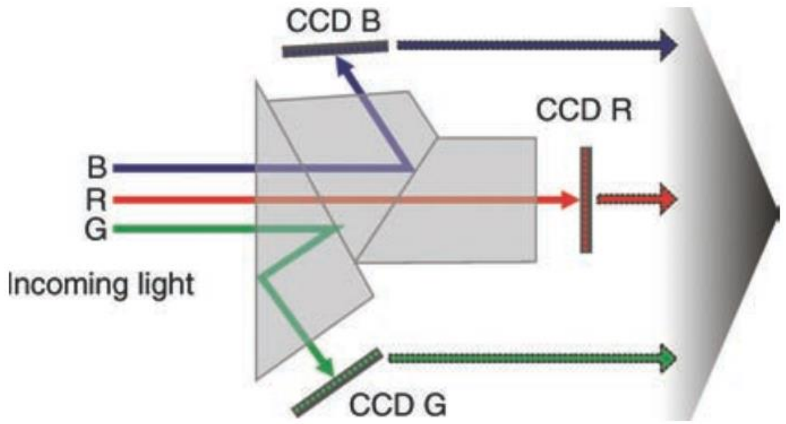


FIGURE 5.7 The beam splitter for three-CCD color cameras.

4.3.1 IMAGE SENSORS

- An alternative technology to CCDs is CMOS. CMOS chips have the advantages of *being cheaper to produce and requiring less power to operate* than comparable CCD chips. Their main disadvantage is the increased susceptibility to noise, which limits their performance at low illumination levels. CMOS sensors were initially used in low-end cameras, such as webcams, but have recently been extended to much more sophisticated cameras, including the Panavision HDMAX 35 mm video camera.
- A representative recent example of CMOS sensors is the Foveon X3 sensor.

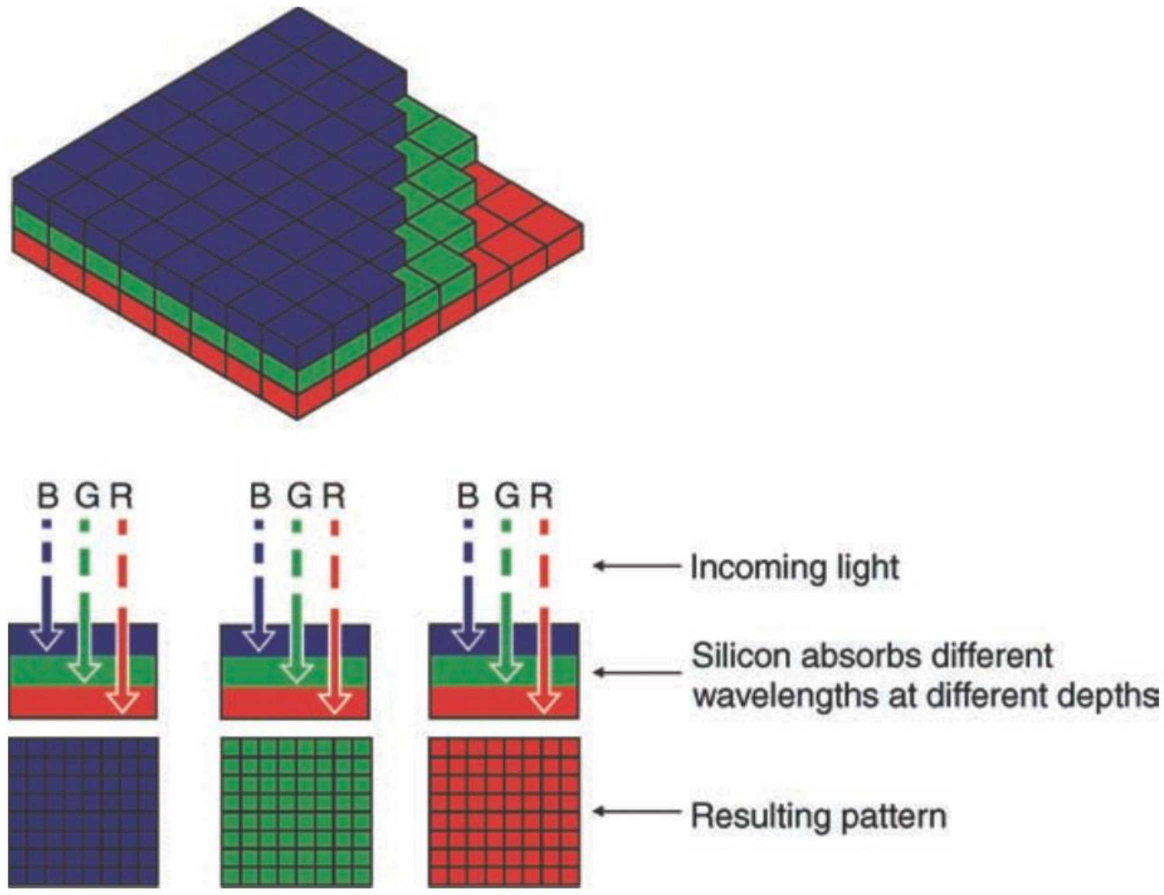


FIGURE 5.8 X3 color sensor.

4.3.2 CAMERA OPTICS

- A camera uses a lens to focus part of the scene onto the image sensor. Two of the most important parameters of a lens are its *magnifying power* and *light gathering capacity*. Magnifying power can be specified by a *magnification factor* (m), which is the ratio between image size and object size:

$$m = \frac{v}{u}$$

  where u is the distance from an object to the lens and v is the distance from the lens to the image plane.
- Lenses may suffer from aberrations, which can affect image quality and generate undesired distortions on the resulting image. Examples of such aberrations include—among many others—the *pincushion distortion* and *the barrel distortion*.

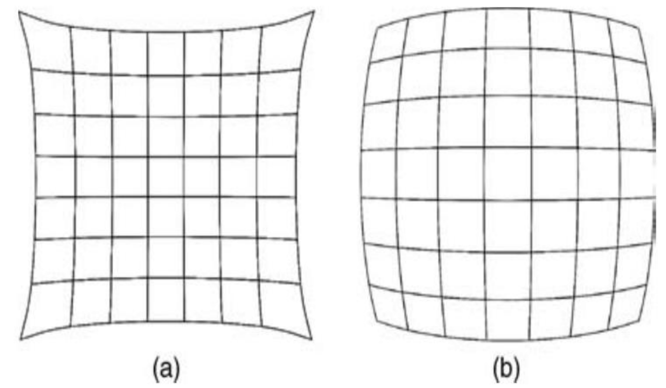


FIGURE 5.10 Examples of lens aberrations: (a) pincushion distortion; (b) barrel distortion.

4.3.2 CAMERA OPTICS

In MATLAB

- The MATLAB Image Acquisition Toolbox (IAT) is a collection of functions that extend the capability of MATLAB, allowing image acquisition operations from a variety of image acquisition devices, from professional-grade frame grabbers to USB-based webcams. The IAT software uses components called *hardware device adaptors* to connect to devices through their drivers (Figure 5.11).

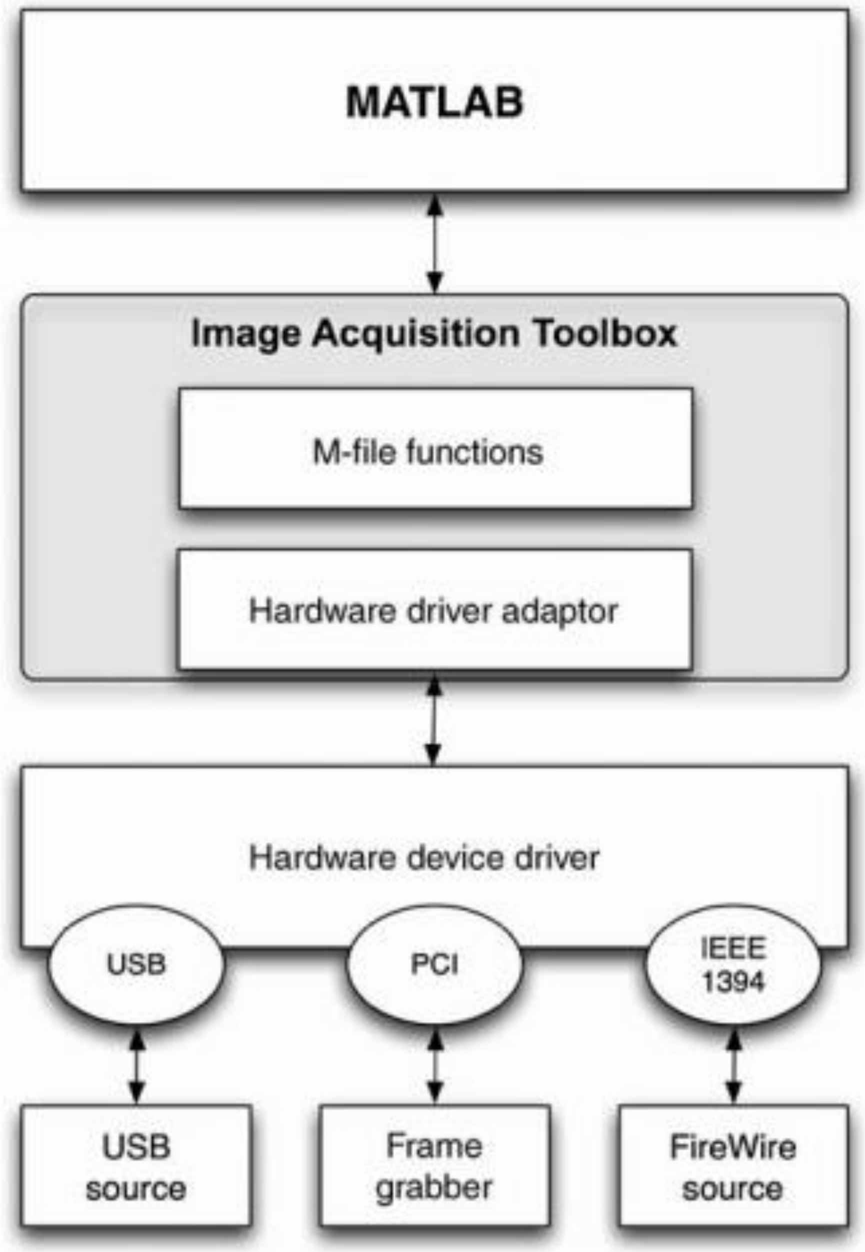


FIGURE 5.11 The main components of the MATLAB Image Acquisition Toolbox.

4.4 IMAGE DIGITIZATION

- The image digitization stage bridges the gap between the analog natural world, from which scenes are acquired, and the digital format expected by computer algorithms in charge of *processing, storing, or transmitting this image*.

- Digitization involves two processes: ***sampling*** (in *time or space*) and ***quantization*** (in *amplitude*). These operations may occur in any sequence, but usually sampling precedes quantization.

4.4 IMAGE DIGITIZATION

- Sampling involves selecting *a finite number of points within an interval*, whereas quantization implies *assigning an amplitude value* (within a finite range of possible values) *to each of those points*.

- The result of the digitization process is *a pixel array*, which is a rectangular matrix of picture elements whose values correspond to their intensities (for monochrome images) or color components (for color images).

4.4 IMAGE DIGITIZATION

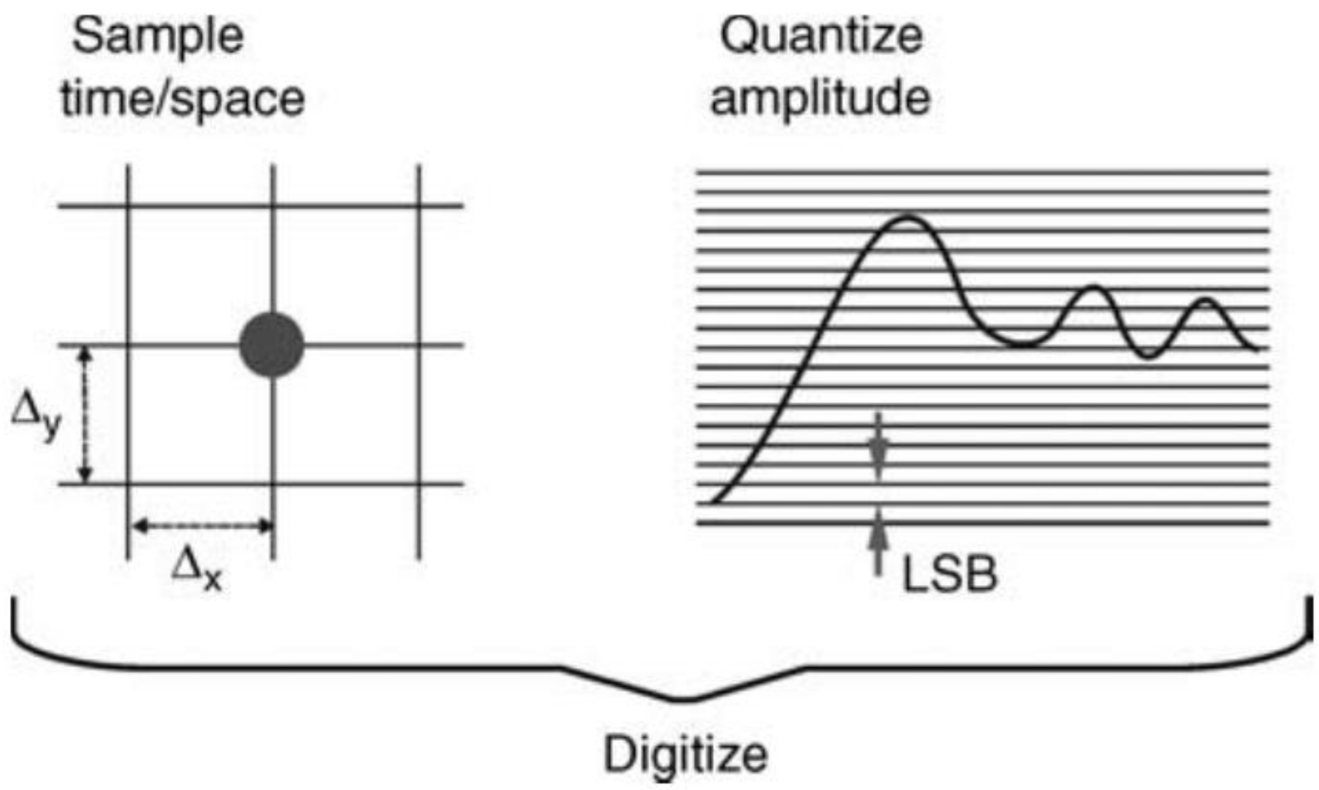


FIGURE 5.12 Digitization = sampling + quantization. Redrawn from [Poy03].

4.4 IMAGE DIGITIZATION

- For consumer cameras and camcorders, it has become common to refer to *the size of the pixel array by the product of the number of pixels* along each dimension and express the result in *megapixels* (Mpx).
- Digitization can take place in many different portions of a machine vision system (MVS) and has been moving progressively closer to the camera hardware during the past few years, making products such as "video capture cards" or "frame grabbers" more of a rarity.

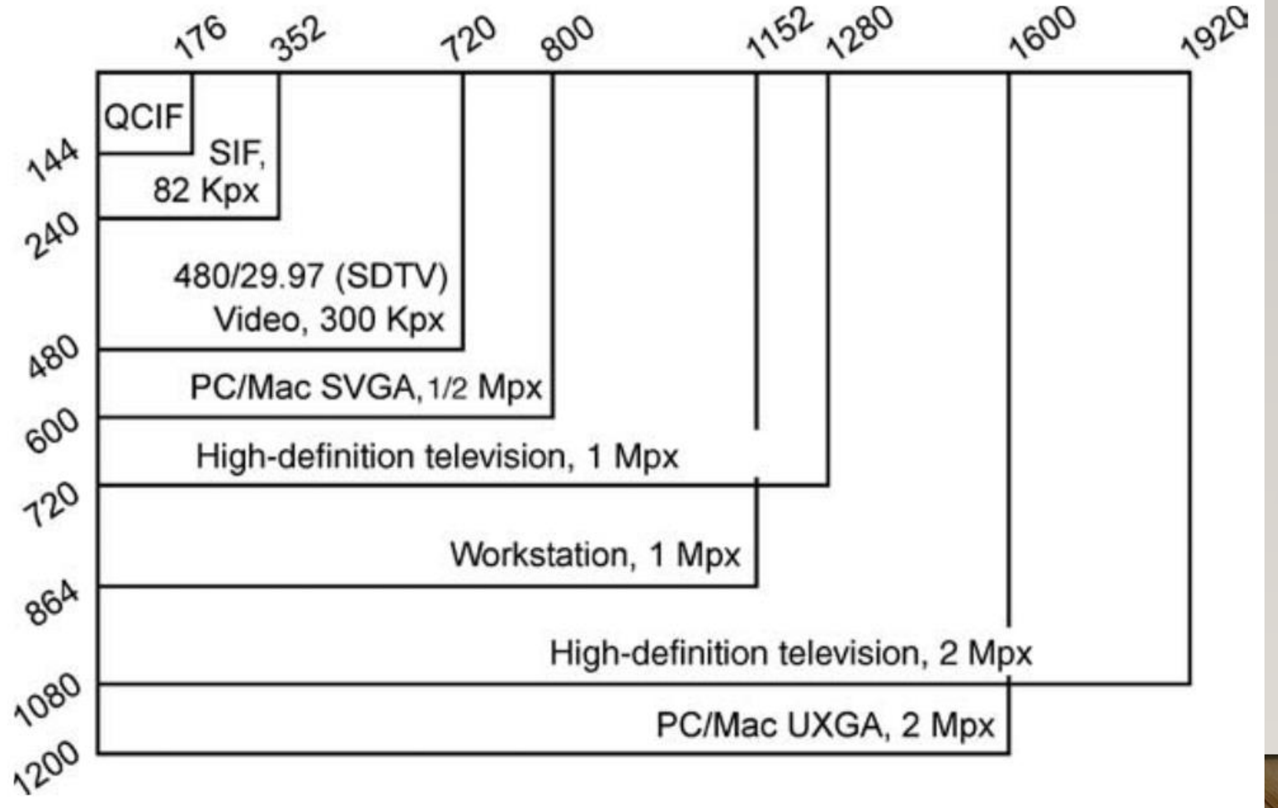


FIGURE 5.13 Pixel arrays of several imaging standards. Redrawn from [Poy03].

4.4.1 SAMPLING

- Sampling is *the process of measuring the value of a 2D function* at discrete intervals along the x and y dimensions.

- A system that has equal horizontal and vertical sampling densities is said to have square sampling. Several imaging and video systems use sampling lattices where the horizontal and the vertical sample pitch are unequal, that is, nonsquare sampling.

4.4.1 SAMPLING

- Two parameters must be taken into account when sampling images:

1. ***The sampling rate***, *that is, the number of samples across the height and width of the image. The choice of an appropriate sampling rate will impact image quality. Inadequate values may lead to a phenomenon known as aliasing.*

2. ***The sampling pattern***, *that is, the physical arrangement of the samples. A rectangular pattern, in which pixels are aligned horizontally and vertically into rows and columns, is by far the most common form, but other arrangements are possible, for example, the hexagonal and log-polar sampling patterns.*

4.4.1 SAMPLING

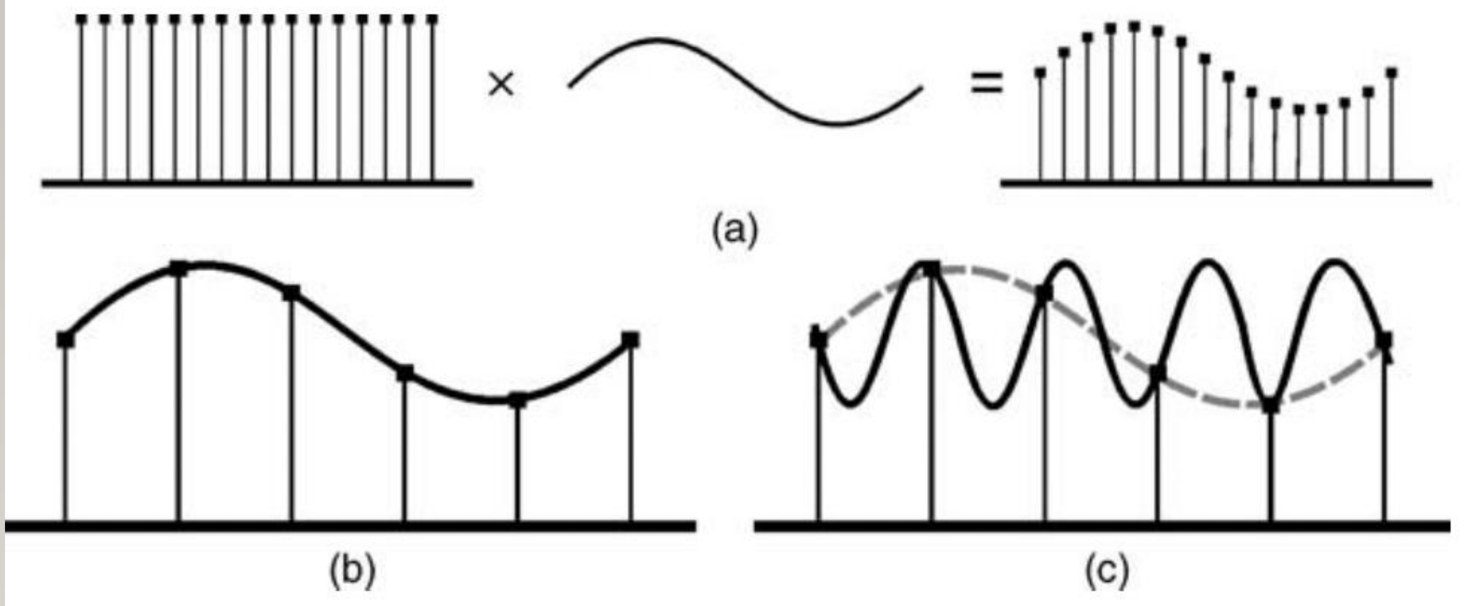


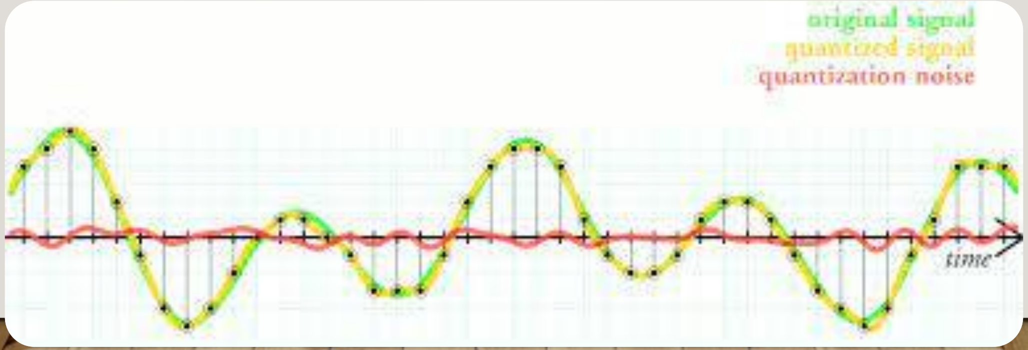
FIGURE 5.14 1D aliasing explanation. Redrawn from [Wat00].

- If sampling takes place at a rate lower than *twice the highest frequency component of the signal* (*the Nyquist criterion*), there will not be enough points to ensure proper reconstruction of the original signal, which is referred to as undersampling or aliasing.

Figure 5.14 illustrates the concept for 1D signals: part (a) shows the sampling process as a product of a train of sampling impulses and the analog signal being sampled, part (b) shows the result of reconstructing a signal from an appropriate number of samples, and part (c) shows that the same number of samples as in (b) would be insufficient to reconstruct a signal with higher frequency.

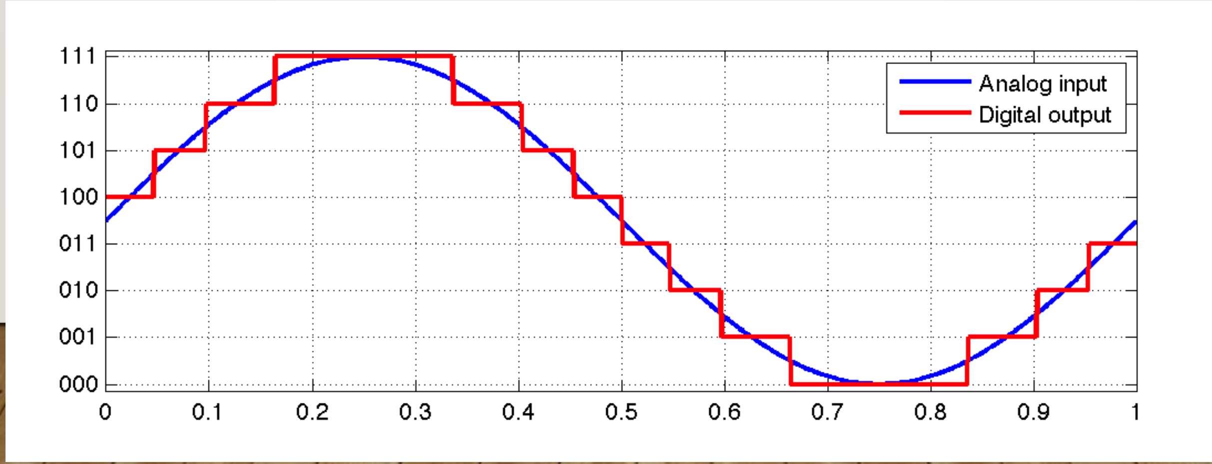
4.4.2 QUANTIZATION

- Quantization is *the process of replacing a continuously varying function* with a discrete set of quantization levels.
- In the case of images, the function is $f(x, y)$ and the quantization levels are also known as *gray levels*. It is common to adopt N quantization levels for image digitization, where N is usually an integral power of 2 that is, $N = 2^n$, where n is the number of bits needed to encode each pixel value.



4.4.2 QUANTIZATION

- The case where $n = 28 = 256$ produces images where each pixel is represented by an unsigned byte, with values ranging from 0 (black) to 255 (white).
- Image quantization can be described as a *mapping process* by which groups of data points (several pixels within a range of gray values) are mapped to a single point (i.e., a single gray level).



4.4.2 QUANTIZATION

- This process is illustrated in Figure 5.15, which depicts the case where the number of gray levels is reduced from 256 to 4 by uniform quantization, meaning that the input gray level range is divided into N equal intervals of length 64.

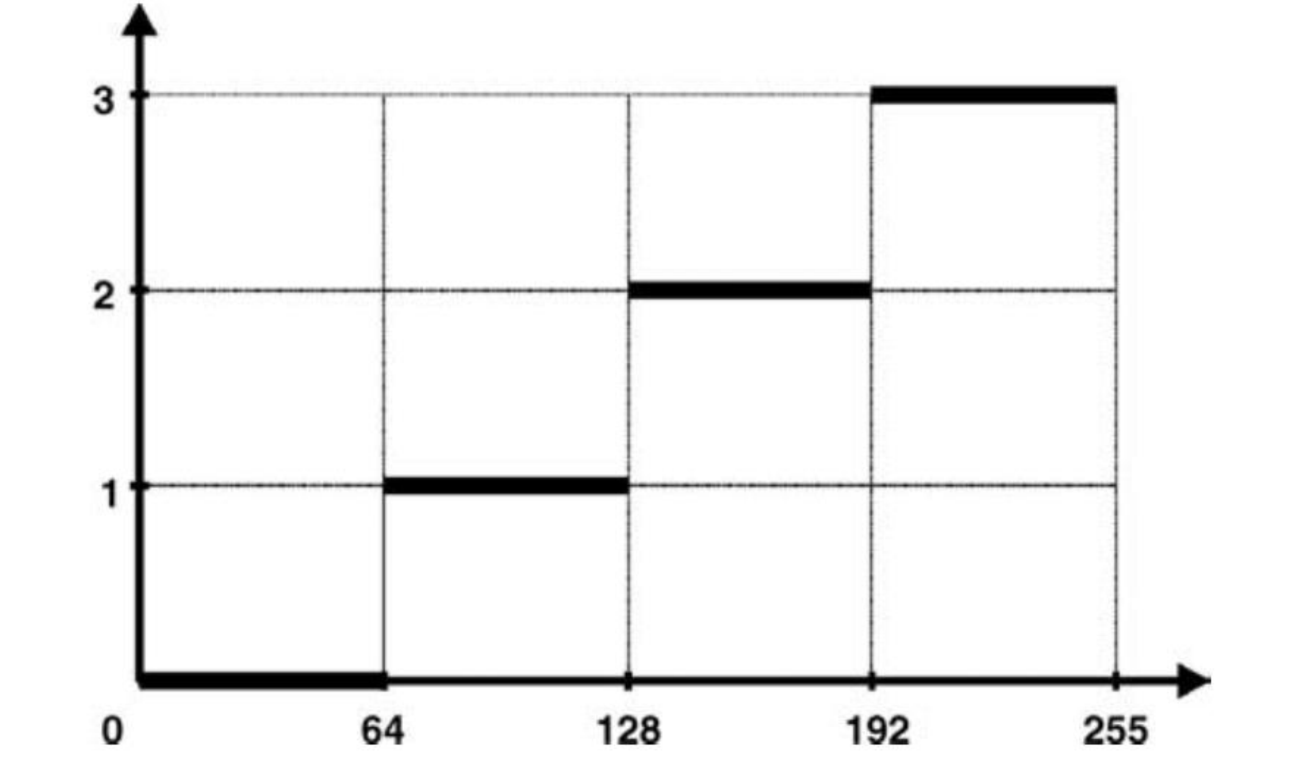


FIGURE 5.15 A mapping function for uniform quantization ($N = 4$).

4.4.3 SPATIAL AND GRAY-LEVEL RESOLUTION

- *Spatial resolution* is a way of expressing *the density of pixels in an image*: the greater the spatial resolution, the more pixels are used to display the image within a certain fixed physical size. It is usually expressed quantitatively using units such as *dots per inch* (*dpi*).
- Gray-level resolution refers to the smallest change in intensity level that the HVS can discern. The adoption of 8 bits per pixel for monochrome images is *a good compromise between subjective quality and practical implementation* (each pixel value is neatly aligned with a byte).
- Higher end imaging applications may require more than 8 bits per color channel and some image file formats support such need (e.g., 12-bit RAW and 16-bit TIFF files).

4.4.3 SPATIAL AND GRAY-LEVEL RESOLUTION

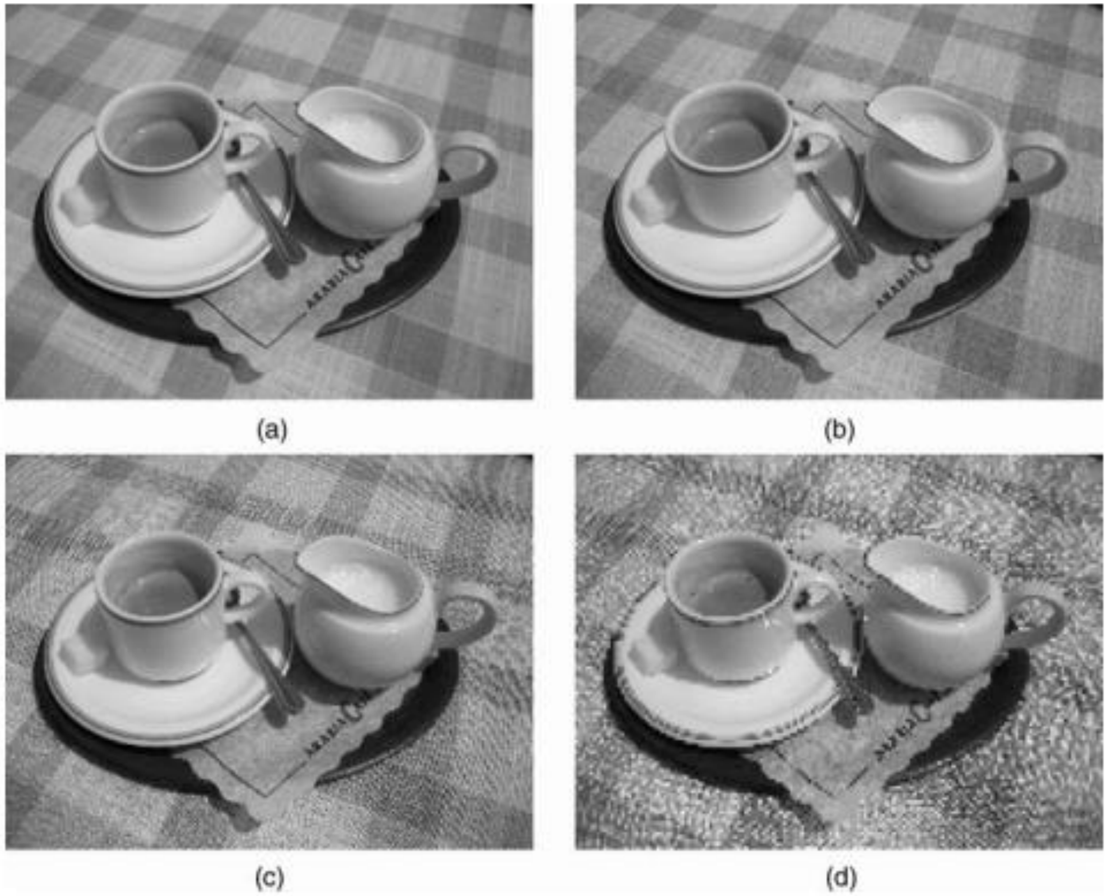


FIGURE 5.16 Effects of sampling resolution on image quality: (a) A 1944 × 2592 image, 256 gray levels, at a 1250 dpi resolution. The same image resampled at (b) 300 dpi; (c) 150 dpi; (d) 72 dpi.

4.4.3 SPATIAL AND GRAY-LEVEL RESOLUTION

- (Re-)quantizing an image in MATLAB can be accomplished using the *grayslice* function.

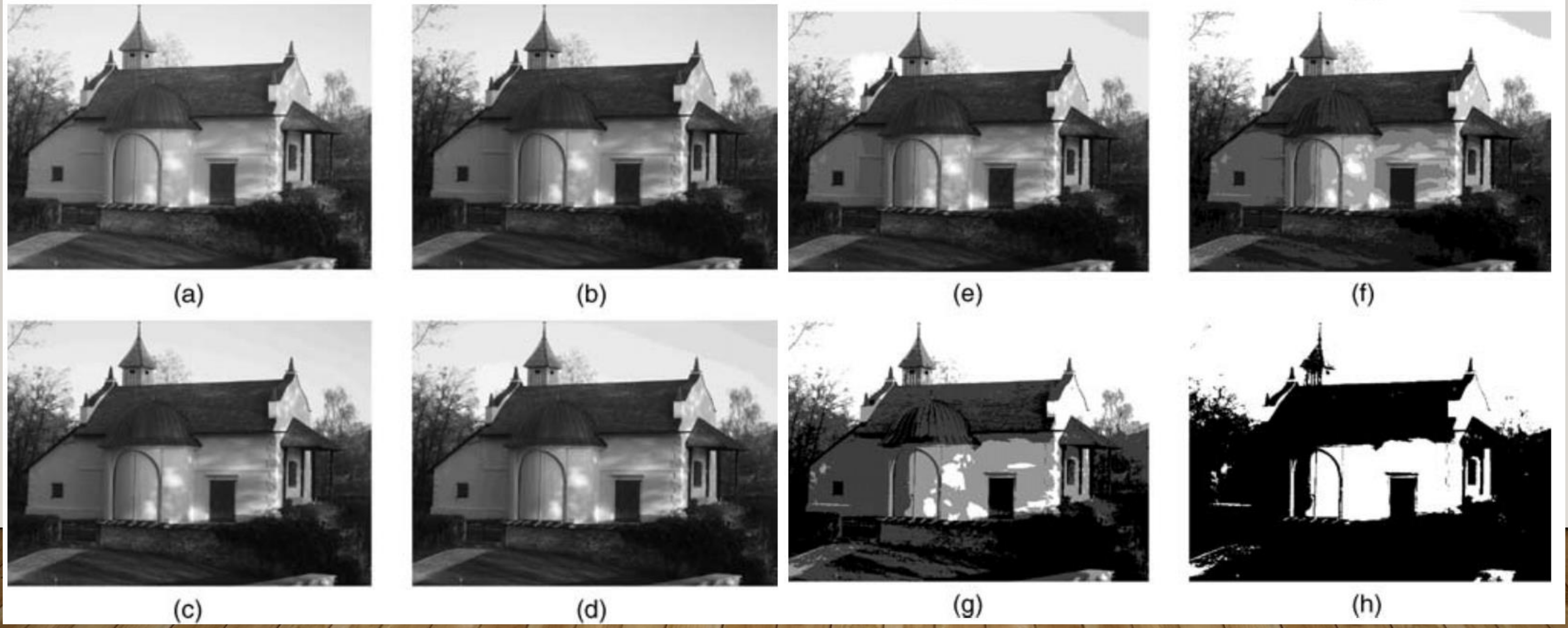


FIGURE 5.17 (a) A 480 x 640 image, 256 gray levels; (b–h) image requantized to 128, 64, 32, 16, 8, 4, and 2 gray levels.

4.4.3 SPATIAL AND GRAY-LEVEL RESOLUTION

MATLAB Code

```
I1 = imread('nature.jpg');
Iplus = im2uint8(rgb2gray(I1));
I2 = grayslice(Iplus,128); figure, imshow(I2,gray(128));
I3 = grayslice(Iplus,64); figure, imshow(I3,gray(64));
I4 = grayslice(Iplus,32); figure, imshow(I4,gray(32));
I5 = grayslice(Iplus,16); figure, imshow(I5,gray(16));
I6 = grayslice(Iplus,8); figure, imshow(I6,gray(8));
I7 = grayslice(Iplus,4); figure, imshow(I7,gray(4));
I8 = grayslice(Iplus,2); figure, imshow(I8,gray(2));
```

WHAT HAVE WE LEARNED?

- Images are formed as *a result of the interaction between the source of radiation* (e.g., visible light), the properties of the objects and surfaces involved, and the relative positioning and properties of the image sensor.
- Images can be classified into *three categories according to the type of interaction* between the source of radiation, the properties of the objects involved, and the relative positioning of the image sensor: *reflection images*, *emission images*, and *absorption images*.

WHAT HAVE WE LEARNED?

- Contemporary image sensors (imagers) are usually built upon CCD or CMOS solid-state technologies. A CCD sensor is made up of an array of light-sensitive cells called "*photosites*," each of which produces a voltage proportional to the intensity of light falling on them. The cells are combined into a (1D or 2D) array that can be *read sequentially by a computer input process*.
- *Image digitization* is the process of sampling a continuous image (in space) and quantizing the resulting amplitude values so that they fall within a finite range.

WHAT HAVE WE LEARNED?

- Some of the main parameters that impact the digitization of an image are *the total number of pixels* and *the maximum number of colors* (or gray levels) per pixel.

- *Sampling* is the process of measuring the value of a function at discrete intervals. (Re-)sampling an image in MATLAB can be accomplished using the *imresize* function.

- *Quantization* is the process of replacing a continuously varying function with a discrete set of quantization levels. (Re-)quantizing an image in MATLAB can be accomplished using the *grayslice* function.

PROBLEMS

Problem 1. Experiment with a scanner's settings (in dpi) and scan the same material (e.g., a photo) several times using different settings but the same file format. Compare the resulting file size and quality for each resulting scanned image.

Problem 2. Repeat Problem 1, but this time keeping the settings the same and changing only the file format used to save the resulting image. Are there significant differences in file size? What about the subjective quality?

PROBLEMS

Problem 3. Assuming a monochrome image with 1024 x 1024 pixels and 256 gray levels,

(a) Calculate the total file size (in bytes), assuming a header (containing basic information, such as width and height of the image) of 32 bytes and no compression.

(b) Suppose the original image has been subsampled by a factor of 2 in both dimensions. Calculate the new file size (in bytes), assuming the header size has not changed.

(c) Suppose the original image has been requantized to allow encoding 2 pixels per byte. Calculate the new file size (in bytes), assuming the header size has not changed.

(d) How many gray levels will the image in part (c) have?

END OF LECTURE 4

LECTURE 5 – ARITHMETIC AND LOGIC, GEOMETRICS OPERATIONS