UNIT I

IMAGE PERCEPTION AND SAMPLING

Elements of Digital Image Processing Systems

Elements of Visual Perception – structure of human eye

– light, luminance, brightness and contrast, mach band effect

– image formation – Monochrome vision model

– Image fidelity criteria – Color vision model –

Introduction to 2D system – linear, shift invarient, casual and stable systems

– optical and modulation transfer function

– 2D sampling theorem – Aliasing

– reconstruction of image from samples

– Practical limitations in sampling and reconstruction

– Image quantization – Lloyd max quantizer – Optimum mean square uniform quantizer.

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image.

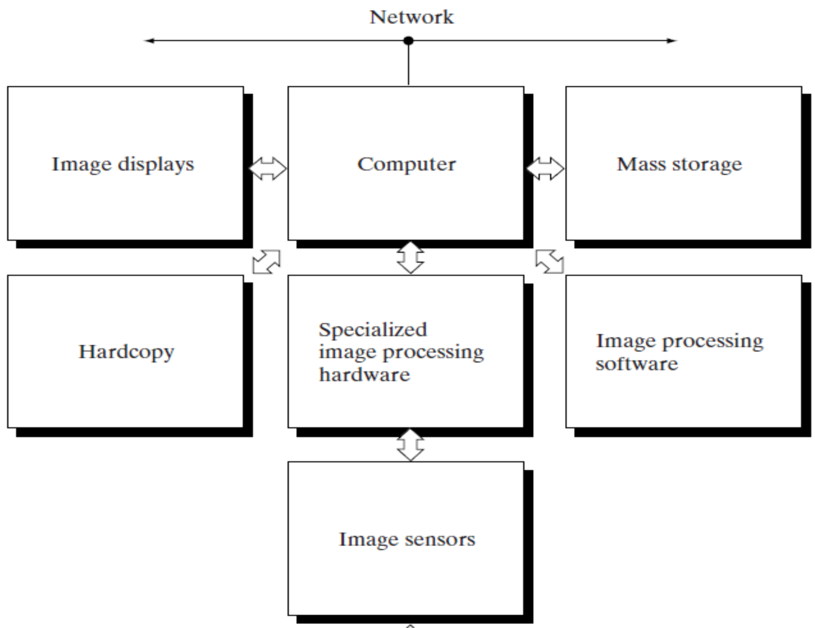
An Image may be defined as a two dimensional function f (x,y) where x & y are spatial (plane) coordinates, and the amplitude of f at any pair of coordinates (x,y) is called intensity or gray level of the image at that point. When x,y and the amplitude values of f are all finite, discrete quantities we call the image as Digital Image.

Applications of Image Processing:

- Intelligent Transportation Systems
- Remote Sensing
- Moving object tracking
- Defense surveillance
- Biomedical Imaging techniques

Elements of Digital Image Processing Systems

- **Image Sensors**
  With reference to sensing, two elements are required to acquire digital image. The first is a physical device that is sensitive to the energy radiated by the object we wish to image and second is specialized image processing hardware.
- **Specialize image processing hardware**
  It consists of the digitizer just mentioned, plus hardware that performs other primitive operations such as an arithmetic logic unit, which performs arithmetic such addition and subtraction and logical operations in parallel on images
- **Computer**
  It is a general purpose computer and can range from a PC to a supercomputer depending on the application. In dedicated applications, sometimes specially designed computer are used to achieve a required level of performance
- **Software**
  It consist of specialized modules that perform specific tasks a well designed package also includes capability for the user to write code, as a minimum, utilizes the specialized module. More sophisticated software packages allow the integration of these modules.



- **Mass storage**
  This capability is a must in image processing applications. An image of size 1024 x1024 pixels ,in which the intensity of each pixel is an 8- bit quantity requires one megabytes of storage space if the image is not compressed .Image processing applications falls into three principal categories of storage
  i) Short term storage for use during processing
  ii) On line storage for relatively fast retrieval
  iii) Archival storage such as magnetic tapes and disks
- **Image displays**
  Image displays in use today are mainly color TV monitors. These monitors are driven by the outputs of image and graphics displays cards that are an integral part of computer system
- **Hardcopy devices**
  The devices for recording image includes laser printers, film cameras, heat sensitive devices inkjet units and digital units such as optical and CD ROM disk. Films provide the highest possible resolution, but paper is the obvious medium of choice for written applications.
- **Networking**
  It is almost a default function in any computer system in use today because of the large amount of data inherent in image processing applications. The key consideration in image transmission bandwidth.

Elements of Visual Perception

Structure of the human Eye

The eye is nearly a sphere with average approximately 20 mm diameter. The eye is enclosed with three membranes

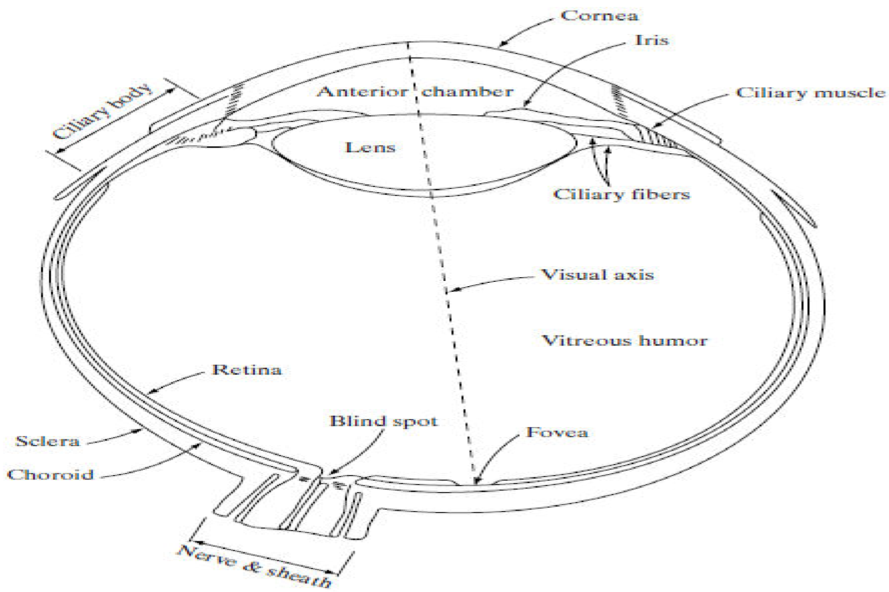
The cornea and sclera - it is a tough, transparent tissue that covers the anterior surface of the eye. Rest of the optic globe is covered by the sclera

The choroid – It contains a network of blood vessels that serve as the major source of nutrition to the eyes. It helps to reduce extraneous light entering in the eye It has two parts

(1) Iris Diaphragms- it contracts or expands to control the amount of light that enters the eyes

(2) Ciliary body

Retina – it is innermost membrane of the eye. When the eye is properly focused, light from an object outside the eye is imaged on the retina. There are various light receptors over the surface of the retina



The two major classes of the receptors are-

1) cones- it is in the number about 6 to 7 million. These are located in the central portion of the retina called the fovea. These are highly sensitive to color. Human can resolve fine details with these cones because each one is connected to its own nerve end. Cone vision is called photopic or bright light vision

2) Rods – these are very much in number from 75 to 150 million and are distributed over the entire retinal surface. The large area of distribution and the fact that several roads are connected to a single nerve give a general overall picture of the field of view. They are not involved in the color vision and are sensitive to low level of illumination. Rod vision is called is scotopic or dim light vision. The absent of reciprocators is called blind spot

Light

The strength of the radiation from a light source is measured using the unit called the candela, or candle power. The total energy from the light source, including heat and all electromagnetic radiation, is called radiance and is usually expressed in watts.

Light is the electromagnetic radiation that stimulates our visual response. It is expressed as a spectral energy distribution $L(\lambda)$, where λ is the wavelength that lies in the visible region, 350 nm to 780 nm, of the electromagnetic spectrum. Light received from an object can be written as

$$I(\lambda) = \rho(\lambda)L(\lambda)$$

where $\rho(\lambda)$ represents the reflectivity or transmissivity of the object and $L(\lambda)$ is the incident energy distribution.

Luminance

Luminance is a measure of the light strength that is actually perceived by the human eye. Radiance is a measure of the total output of the source; luminance measures just the portion that is perceived.

The *luminance* or *intensity* of a spatially distributed object with light distribution $I(x, y, \lambda)$ is defined as

$$f(x, y) = \int_0^\infty I(x, y, \lambda)V(\lambda)\, d\lambda$$

where $V(\lambda)$ is called the *relative luminous efficiency function* of the visual system.

Brightness

The brightness or apparent brightness of an object is the perceived luminance and depends on the luminance of the surround. Brightness is a subjective, psychological measure of perceived intensity.Brightness is practically impossible to measure objectively. It is relative. For example, a burning candle in a darkened room will appear bright to the viewer; it will not appear bright in full sunshine.

Contrast

simultaneous contrast is related to the fact that a region's perceived brightness does not depend simply on its intensity, as below fig. demonstrates.All the center squares have exactly the same intensity.

Contrast is defined as the ratio (max-min)/(max+min)

where max and min are the maximum and minimum of the grating intensity,respectively.

FIGURE Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

Mach band effect

We know that perceived brightness is not a simple function of intensity. The visual system tends to undershoot or overshoot around the boundary of regions of different intensities. Figure shows a striking example of this phenomenon. Although the intensity of the stripes is constant, we actually perceive a brightnesspattern that is strongly scalloped, especially near the boundaries. These seemingly scalloped bands are called *Mach bands (* Ernst Mach, 1865).

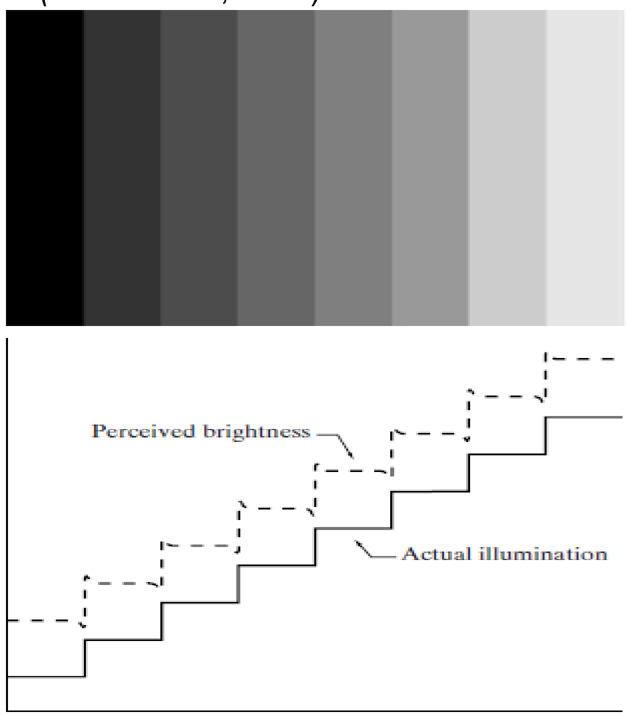
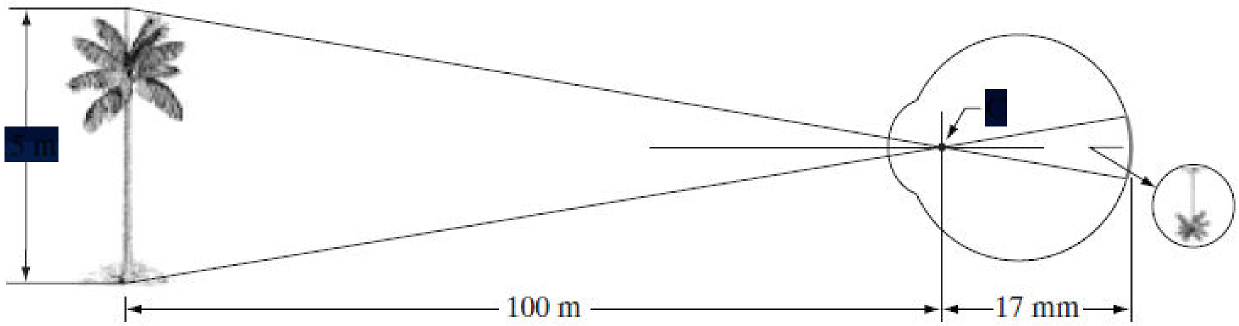


Fig. (a) An example showing that perceived brightness is not a simple function of intensity. The relative vertical positions between the two profiles in (b) have no special significance; they were chosen for clarity.

Image Formation

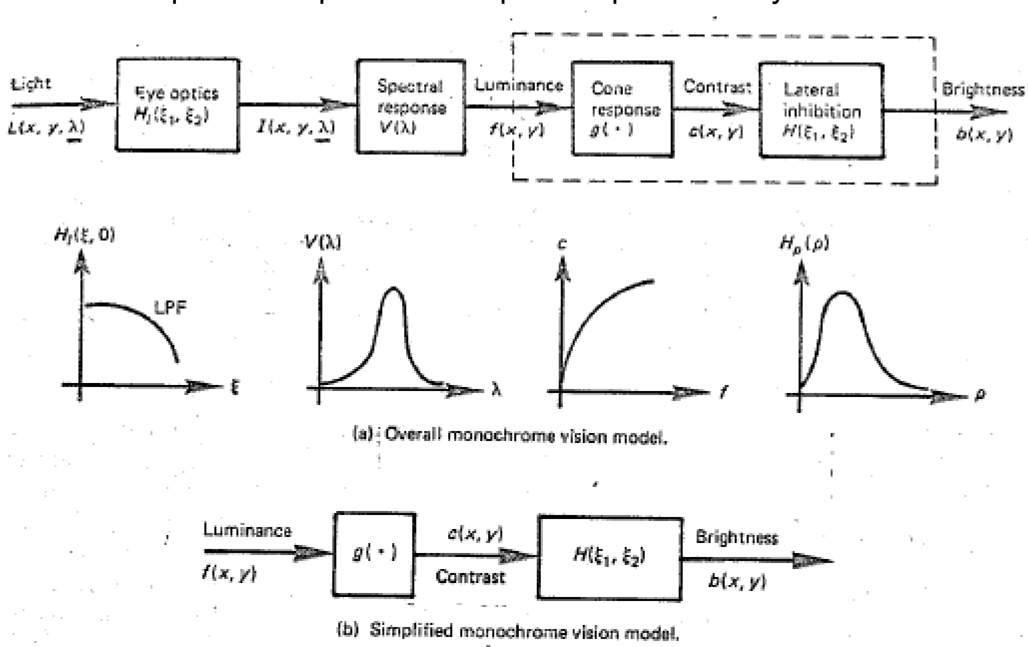
The principal difference between the lens of the eye and an ordinary optical lens is that the former is flexible. The radius of curvature of the anterior surface of the lens is greater than the radius of its posterior surface. The shape of the lens is controlled by tension in the fibers of the ciliary body. To focus on distant objects, the controlling muscles cause the lens to be relatively flattened. Similarly, these muscles allow the lens to become thicker in order to focus on objects near the eye.



The distance between the center of the lens and the retina (called the *focal length*) varies from approximately 17 mm to about 14 mm, as the refractive power of the lens increases from its minimum to its maximum. When the eye focuses on an object farther away than about 3 m, the lens exhibits its lowest refractive power. When the eye focuses on a nearby object, the lens is most strongly refractive. This information makes it easy to calculate the size of the retinal image of any object. The retinal image is reflected primarily in the area of the fovea. Perception then takes place by the relative excitation of light receptors, which transform radiant energy into electrical impulses that are ultimately decoded by the brain.

Monochrome vision model

- The logarithmic/linear system eye model provides a reasonable prediction of visual response over a wide range of intensities.
- However, at high spatial frequencies and at very low or very high intensities, observed responses depart from responses predicted by the model.



(a) Overall monochrome vision model.

(b) Simplified monochrome vision model.

Image fidelity criteria

Image fidelity criteria are useful for measuring image quality and for rating the performance of a processing technique or a vision system. There are two types of criteria that are used for evaluation of image quality, subjective and quantitative. The subjective criteria use rating scales such as goodness scales and impairment scales. A goodness scale may be a global scale or a group scale (Table 3.2). The overall goodness criterion rates image quality on a scale ranging from excellent to unsatisfactory. A training set of images is used to calibrate such a scale. The group goodness scale is based on comparisons within a set of images.

The impairment scale (Table 3.3) rates an image on the basis of the level of degradation present in an image when compared with an *ideal* image. It is useful in applications such as image coding, where the encoding process introduces degradations in the output image.

TABLE 3.2 Image Goodness Scales

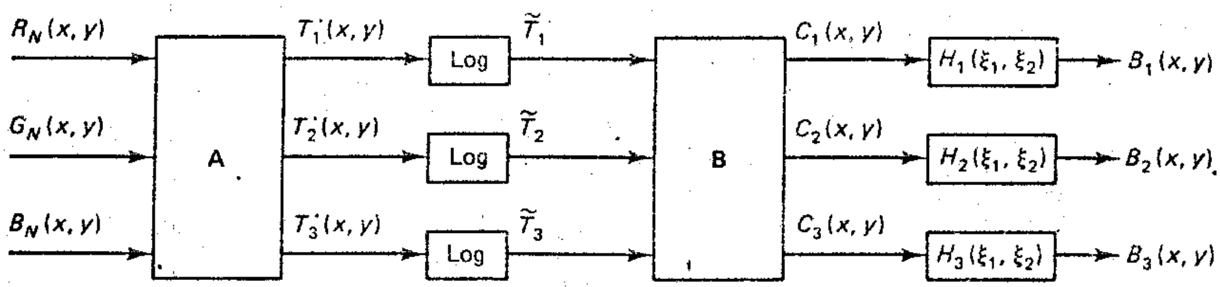
Overall goodness scale		Group goodness scale	
Excellent	(5)	Best	(7)
Good	(4)	Well above average	(6)
Fair	(3)	Slightly above average	(5)
Poor	(2)	Average	(4)
Unsatisfactory	(1)	Slightly below average	(3)
		Well below average	(2)
		Worst	(1)

The numbers in parenthesis indicate a numerical weight attached to the rating.

TABLE 3.3 Impairment Scale

Not noticeable	(1)
Just noticeable	(2)
Definitely noticeable but only slight impairment	(3)
Impairment not objectionable	(4)
Somewhat objectionable	(5)
Definitely objectionable	(6)
Extremely objectionable	(7)

Color vision model



Introduction to 2D system

A one-dimensional continuous signal will be represented as a function of one variable: $f(x)$, $u(x)$, $s(t)$, and so on. One-dimensional sampled signals will be written as single index sequences: u_n, $u(n)$, and the like.

A continuous image will be represented as a function of two independent variables: $u(x, y)$, $v(x, y)$, $f(x, y)$, and so forth. A sampled image will be represented as a two- (or higher) dimensional sequence of real numbers: $u_{m,n}$, $v(m, n)$, $u(i, j, k)$,

Linear, shift invarient, casual and stable systems

A large number of imaging systems can be modeled as two-dimensional linear systems. Let $x(m, n)$ and $y(m, n)$ represent the input and output sequences, respectively, of a two dimensional system

$$y(m, n) = \mathcal{H}[x(m, n)]$$

This system is called *linear* if and only if any linear combination of two inputs $x_1(m, n)$ and $x_2(m, n)$ produces the same combination of their respective outputs $y_1(m, n)$ and $y_2(m, n)$, i.e., for arbitrary constants a_1 and a_2

A system is called shift invariant if translation of the input causes translation of the output. For shift invariant system, the output is

$$y(m, n) = \sum\sum_{m', n' = -\infty}^{\infty} h(m - m', n - n')x(m', n')$$

A one-dimensional shift invariant system is called causal if its output at any time is not affected by future inputs. This means its impulse response $h(n) = 0$ for $n < 0$ and its transfer function must have a one-sided Laurent series, i.e.,

$$H(z) = \sum_{n=0}^{\infty} h(n) z^{-n}$$

A system is called *stable* if its output remains uniformly bounded for any bounded input. For linear shift invariant systems, this condition requires that the impulse response should be absolutely summable

$$\sum_{n=-\infty}^{\infty} |h(n)| < \infty$$

Optical and modulation transfer function

For a spatially invariant imaging system, its optical transfer function (OTF) is defined as its normalized frequency response, that is

$$\text{OTF} = \frac{H(\xi_1, \xi_2)}{H(0,0)}$$

The modulation transfer function (MTF) is defined as the magnitude of OTF, that is

$$\text{MTF} = |\text{OTF}| = \frac{|H(\xi_1, \xi_2)|}{|H(0,0)|}$$

2D sampling theorem and Aliasing

Sampling means digitizing the co-ordinate value (x, y). Quantization means digitizing the amplitude value.

The Shannon sampling theorem tells us that, if the function is sampled at a rate equal to or greater than twice its highest frequency, it is possible to recover completely the original function from its samples.

If the function is *undersampled*, then a phenomenon called *aliasing* corrupts the sampled image.The corruption is in the form of additional frequency components being introduced into the sampled function. These are called *aliased frequencies*. Note that the *sampling rate* in images is the number of samples taken (in both spatial directions) per unit distance.

Reconstruction of image from samples

From uniqueness of the Fourier tranform, we know that if the spectrum of the original image could be recovered somehow from the spectrum of the sampled image, then we would have the interpolated continuous image from the sampled Image. If the x, y sampling frequencies are greater than the twice the bandwidths,i.e.,

$$\xi_{xs} > 2\xi_{x0}, \qquad \xi_{ys} > 2\xi_{y0}$$

Or, equivalently, if the sampling intervals are smaller than one half of reciprocal of bandwidths,i.e.,

$$\Delta x < \frac{1}{2\xi_{x0}}, \qquad \Delta y < \frac{1}{2\xi_{y0}}$$

Then the original image can be recovered by a low-pass filter with the frequency response. That is the original continuous signal can be recovered exactly by low pass filtering the sampled image.

Practical limitations in sampling and reconstruction

The foregoing sampling theory is based on several idealizations. Real-world images are not bandlimited, which means aliasing errors occur. These can be reduced by low-pass filtering the input image prior to sampling but at the cost of attenuating higher spatial frequencies. Such resolution loss, which results in blurring of the image, also occurs because practical scanners have finite apertures. Finally, the reconstruction system can never be the ideal low-pass filter required by the sampling theory. Its transfer function depends on the display aperture. Figure 4.10 represents the practical sampling/reconstruction systems.

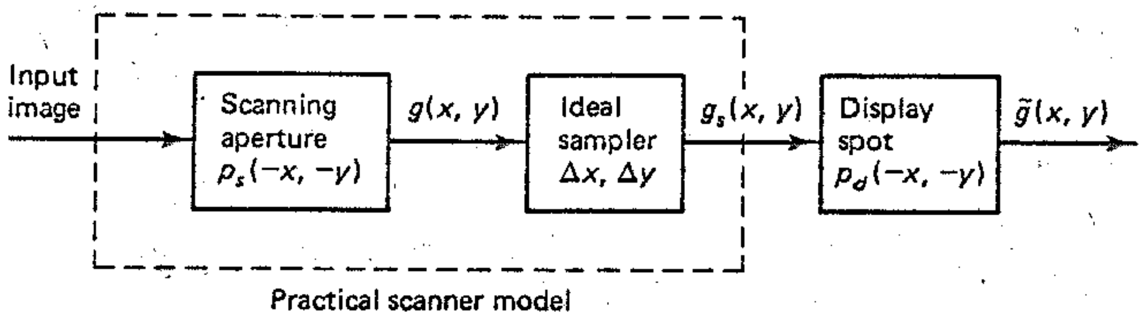
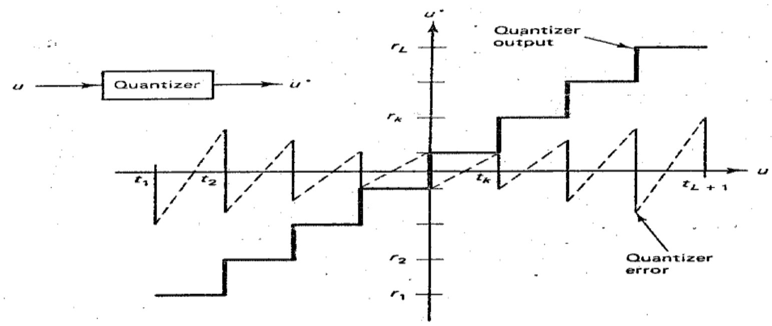


Figure 4.10 Practical sampling and reconstruction. In the ideal case $p_s(x, y) = p_d(x, y) = \delta(x, y)$.

Image quantization

The step subsequent to sampling in image digitization is quantization. A quantizer maps a continuous variable u into a discrete variable $u^\cdot$, which takes values from a finite set $\{r_1, \ldots, r_L\}$ of numbers. This mapping is generally a staircase function (Fig. 4.16) and the quantization rule is as follows: Define $\{t_k, k = 1, \ldots, L+1\}$ as a set of increasing *transition* or *decision levels* with t_1 and t_{L+1} as the minimum and maximum values, respectively, of u. If u lies in interval $[t_k, t_{k+1})$, then it is mapped to r_k, the kth *reconstruction level.*



Lloyd max quantizer – Optimum mean square uniform quantizer

This quantizer minimizes the mean square error for a given number of quantization levels. Let u be a real scalar random variable with a continuous probability density function $p_u(u)$. It is desired to find the decision levels t_k and the reconstruction levels r_k for an L-level quantizer such that the mean square error

Properties of Lloyd max quantizer – Optimum mean square uniform quantizer

This quantizer has the following properties

1. The quantizer output is an unbiased estimate of the input
2. The quantizer error is orthogonal to the quantizer output

Questions for Practice:-

TWO MARKS:

1. What is meant by luminance?
2. Define contrast.
3. What is brightness?
4. Write notes on mach band effect.
5. Write notes on Image fidelity criteria.
6. Write notes on optical and modulation transfer function.
7. Write down 2D sampling theorem.
8. What is meant by Aliasing?
9. Write notes on image quantization
10. Write notes on Optimum mean square uniform quantizer

12 Marks

1. Explain the various Elements of Digital Image Processing Systems
2. Explain the Elements of Visual Perception.
3. Draw and Explain Monochrome vision model and Color vision model.
4. Briefly explain the reconstruction of image from samples and also note on Practical limitations in sampling and reconstruction