IMAGE FUNDAMENTALS

Digital Image Processing

- Digital Image

  a two-dimensional function

  x and *y* are spatial coordinates

  The amplitude of *f* is called intensity or gray level at the point (x, y)
- Digital Image Processing

  Process digital images by means of computer, it covers low-, mid-, and high-level processes.

  low-level: inputs and outputs are images

  mid-level: outputs are the attributes extracted from input images

  high-level: an ensemble of recognition of individual objects
- Pixel : the elements of a digital image

Sources for the images are Electromagnetic (EM) energy spectrum, Acoustic, Ultrasonic, Electronic and Synthetic images produced by computer.

Major uses

Gamma-ray imaging: nuclear medicine and astronomical observations
X-rays: medical diagnostics, industry, and astronomy, etc.
Ultraviolet: lithography, industrial inspection, microscopy, lasers, biological imaging, and astronomical observations
Visible and infrared bands: light microscopy, astronomy, remote sensing, industry, and law enforcement
Microwave band: radar
Radio band: medicine (such as MRI) and astronomy

Fundamental Steps in Digital Image Processing:

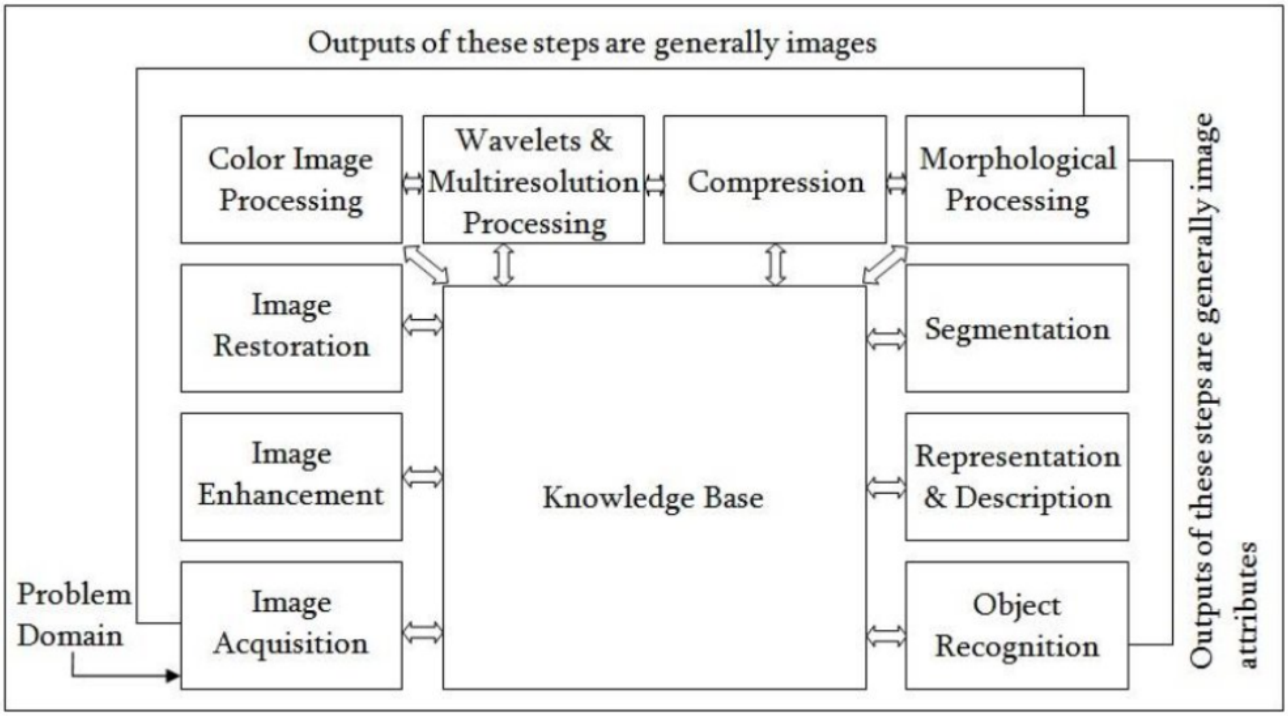
There are some fundamental steps but as they are fundamental, all these steps may have sub-steps. The fundamental steps are described below with a neat diagram.

Image Acquisition: This is the first step or process of the fundamental steps of digital image processing. Image acquisition could be as simple as being given an image that is already in digital form. Generally, the image acquisition stage involves preprocessing, such as scaling etc.

Image Enhancement: Image enhancement is among the simplest and most appealing areas of digital image processing. Basically, the idea behind enhancement techniques is to bring out detail that is obscured, or simply to highlight certain features of interest in an image. Such as, changing brightness & contrast etc.

Image Restoration: Image restoration is an area that also deals with improving the appearance of an image. However, unlike enhancement, which is subjective, image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image degradation.

Color Image Processing: Color image processing is an area that has been gaining its importance because of the significant increase in the use of digital images over the Internet. This may include color modeling and processing in a digital domain etc.



Wavelets and Multiresolution Processing: Wavelets are the foundation for representing images in various degrees of resolution. Images subdivision successively into smaller regions for data compression and for pyramidal representation.

Compression: Compression deals with techniques for reducing the storage required to save an image or the bandwidth to transmit it. Particularly in the uses of internet it is very much necessary to compress data.

Morphological Processing: Morphological processing deals with tools for extracting image components that are useful in the representation and description of shape.

Segmentation: Segmentation procedures partition an image into its constituent parts or objects. In general, autonomous segmentation is one of the most difficult tasks in digital image processing. A rugged segmentation procedure brings the process a long way toward successful solution of imaging problems that require objects to be identified individually.

Representation and Description: Representation and description almost always follow the output of a segmentation stage, which usually is raw pixel data, constituting either the boundary of a region or all the points in the region itself. Choosing a representation is only part of the solution for transforming raw data into a form suitable for subsequent computer processing. Description deals with extracting attributes that result in some quantitative information of interest or are basic for differentiating one class of objects from another.

Object recognition: Recognition is the process that assigns a label, such as, "vehicle" to an object based on its descriptors.

Knowledge Base: Knowledge may be as simple as detailing regions of an image where the information of interest is known to be located, thus limiting the search that has to be conducted in seeking that information. The knowledge base also can be quite complex, such as an interrelated list of all major possible defects in a materials inspection problem or an image database containing high-resolution satellite images of a region in connection with change-detection applications.

A Simple Image Formation Model:

$f(x,y) = i(x,y)\,r(x,y)$

$f(x,y)$: intensity at the point (x,y)

$i(x,y)$: illumination at the point (x,y)

(the amount of source illumination incident on the scene)

$r(x,y)$: reflectance/transmissivity at the point (x,y)

(the amount of illumination reflected/transmitted by the object)

where $0 < i(x,y) < \infty$ and $0 < r(x,y) < 1$

Some Typical Ranges of illumination

Illumination

Lumen — A unit of light flow or luminous flux

Lumen per square meter (lm/m^2) — The metric unit of measure for illuminance of a surface

- On a clear day, the sun may produce in excess of 90,000 lm/m^2 of illumination on the surface of the Earth
- On a cloudy day, the sun may produce less than 10,000 lm/m^2 of illumination on the surface of the Earth
- On a clear evening, the moon yields about 0.1 lm/m^2 of illumination

The typical illumination level in a commercial office is about 1000 lm/m^2

Some Typical Ranges of Reflectance

Reflectance

- 0.01 for black velvet
- 0.65 for stainless steel
- 0.80 for flat-white wall paint
- 0.90 for silver-plated metal and 0.93 for snow

Representing Digital Images:

The representation of an M×N numerical array as

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \cdots & \cdots & \cdots & \cdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}$$

The representation of an M×N numerical array in MATLAB

$$f(x,y) = \begin{bmatrix} f(1,1) & f(1,2) & \cdots & f(1,N) \\ f(2,1) & f(2,2) & \cdots & f(2,N) \\ \cdots & \cdots & \cdots & \cdots \\ f(M,1) & f(M,2) & \cdots & f(M,N) \end{bmatrix}$$

Discrete intensity interval [0, L-1], $L=2^k$
The number b of bits required to store a $M \times N$ digitized image
$b = M \times N \times k$

Spatial and Intensity Resolution

Spatial resolution: The spatial resolution of an image is determined by how sampling was carried out. Spatial resolution is a measure of the smallest discernible detail in an image. It is stated with line pairs per unit distance, dots (pixels) per unit distance, and a dot per inch (dpi) .DPI (dots per inch) is used to measure the spatial resolution.

Intensity resolution: Intensity resolution refers to the smallest discernible change in intensity level. The more intensity levels used, the finer the level of detail discernible in an image. The number of bits used to quantize intensity is often referred as the intensity resolution. It is stated with 8 bits, 12 bits, 16 bits, etc.

Image interpolation: It is a basic tool used in tasks such as zooming, shrinking, rotating, and geometric corrections. Traditional methods include nearest neighbor, bilinear, and bicubic

.

Basic Relationships between Pixels:

- Neighborhood
- Adjacency
- Connectivity
- Paths
- Regions and boundaries

Neighbors of a pixel p at coordinates (x, y)

- **4-neighbors of p**, denoted by $\mathbf{N_4(p)}$:

(x-1, y), (x+1, y), (x,y-1), and (x, y+1).

- **4 diagonal neighbors of p**, denoted by $\mathbf{N_D(p)}$:

(x-1, y-1), (x+1, y+1), (x+1,y-1), and (x-1, y+1).

- **8 neighbors of p**, denoted $\mathbf{N_8(p)}$

$N_8(p) = N_4(p) \cup N_D(p)$

Adjacency

Let V be the set of intensity values

- **4-adjacency**: Two pixels p and q with values from V are 4-adjacent if q is in the set $N_4(p)$.
- **8-adjacency**: Two pixels p and q with values from V are 8-adjacent if q is in the set $N_8(p)$.
- **m-adjacency**: Two pixels p and q with values from V are m-adjacent if
  (i) q is in the set $N_4(p)$, or
  (ii) q is in the set $N_D(p)$ and the set $N_4(p) \cap N_4(p)$ has no pixels whose values are from V.

► **Path**

A (digital) path (or curve) from pixel p with coordinates (x_0, y_0) to pixel q with coordinates (x_n, y_n) is a sequence of distinct pixels with coordinates

$(\mathbf{x_0, y_0}), (\mathbf{x_1, y_1}), \ldots, (\mathbf{x_n, y_n})$

Where (x_i, y_i) and (x_{i-1}, y_{i-1}) are adjacent for $1 \le i \le n$.

- Here *n* is the *length* of the path.
- If $(x_0, y_0) = (x_n, y_n)$, the path is ***closed*** path.
- We can define 4-, 8-, and m-paths based on the type of adjacency used.

Connectivity and regions

Let S represent a subset of pixels in an image

- For every pixel *p* in S, the set of pixels in S that are connected to *p* is called a *connected component* of S.
- If S has only one connected component, then S is called *Connected Set*.
- We call R a region of the image if R is a connected set
- Two regions, R_i and R_j are said to be *adjacent* if their union forms a connected set.
- Regions that are not to be adjacent are said to be *disjoint.*

Boundary (or border)

- The *boundary* of the region R is the set of pixels in the region that have one or more neighbors that are not in R.
- If R happens to be an entire image, then its boundary is defined as the set of pixels in the first and last rows and columns of the image.

Distance Measures

Given pixels *p, q* and *z* with coordinates (x, y), (s, t), (u, v) respectively, the distance function D has following properties:

a. $D(p, q) \ge 0$ [$D(p, q) = 0$, iff $p = q$]
b. $D(p, q) = D(q, p)$
c. $D(p, z) \le D(p, q) + D(q, z)$

The following are the different Distance measures:

a. Euclidean Distance: $D_e(p, q) = [(x-s)^2 + (y-t)^2]^{1/2}$
b. City Block Distance: $D_4(p, q) = |x-s| + |y-t|$
c. Chess Board Distance: $D_8(p, q) = \max(|x-s|, |y-t|)$

Image perception:

Structure of the human eye

Three membranes enclose the eye: the cornea and sclera, choroid, and retina. At its anterior extreme, the choroid is divided into the ciliary body and the iris; the later contracts or expands to control the amount of light that enters the eye. When the eye is properly focused, light from an object outside the eye is imaged on the retina

• Two kinds of light receptors distribute on the retina, cones and rods

• Cones are primarily located in the central portion of the retina, called fovea and are sensitive to color; they function best in relatively bright light; so, cone vision is called bright-light vision

• Rods are distributed over the retinal surface; rods serve to give a general overall picture of the field of view; they are not involved in color vision and are sensitive to low levels of illumination; rod vision is called dim-light vision.
• Around the region of the emergence of the optic nerve, there is no receptors and results in the so-called blind spot.
Distribution of rods and cones in retina

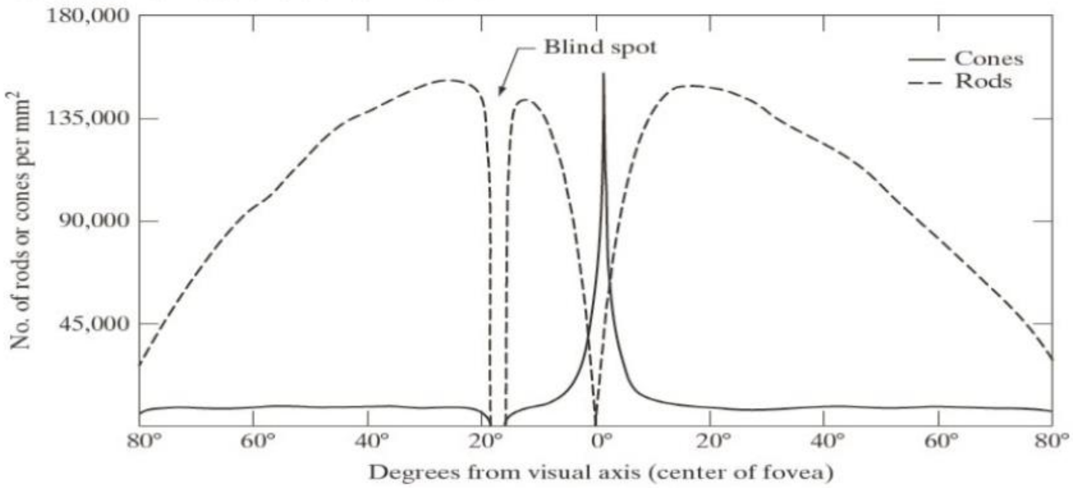
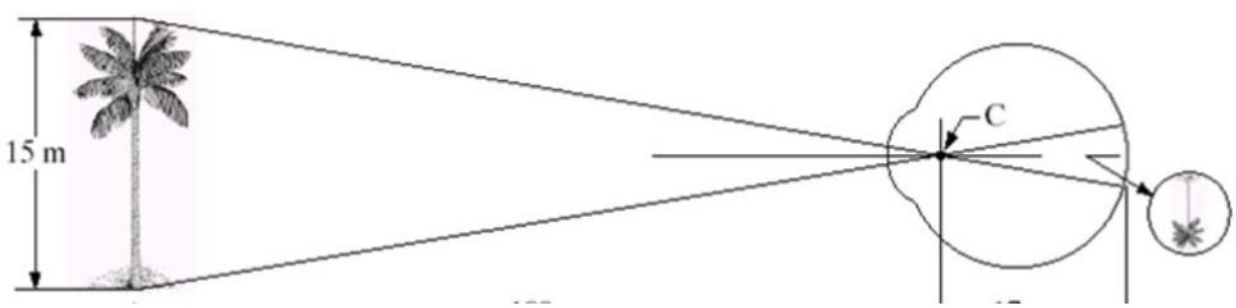


Image formation in the eye

Muscles within the eye can be used to change the shape of the lens allowing us focus on objects that are near or far away. An image is focused onto the retina causing rods and cones to become excited which ultimately send signals to the brain.



The principal difference between the lens of the eye and an ordinary optical lens is that the former is flexible. As illustrated in Fig. 2.1, 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.

Image Sampling and Quantization:

Sampling and quantization will convert a continuous image signal *f* to a discrete digital form.

- Sampling: digitizing the 2-dimensional spatial coordinate values
- Quantization: digitizing the amplitude values (brightness level)

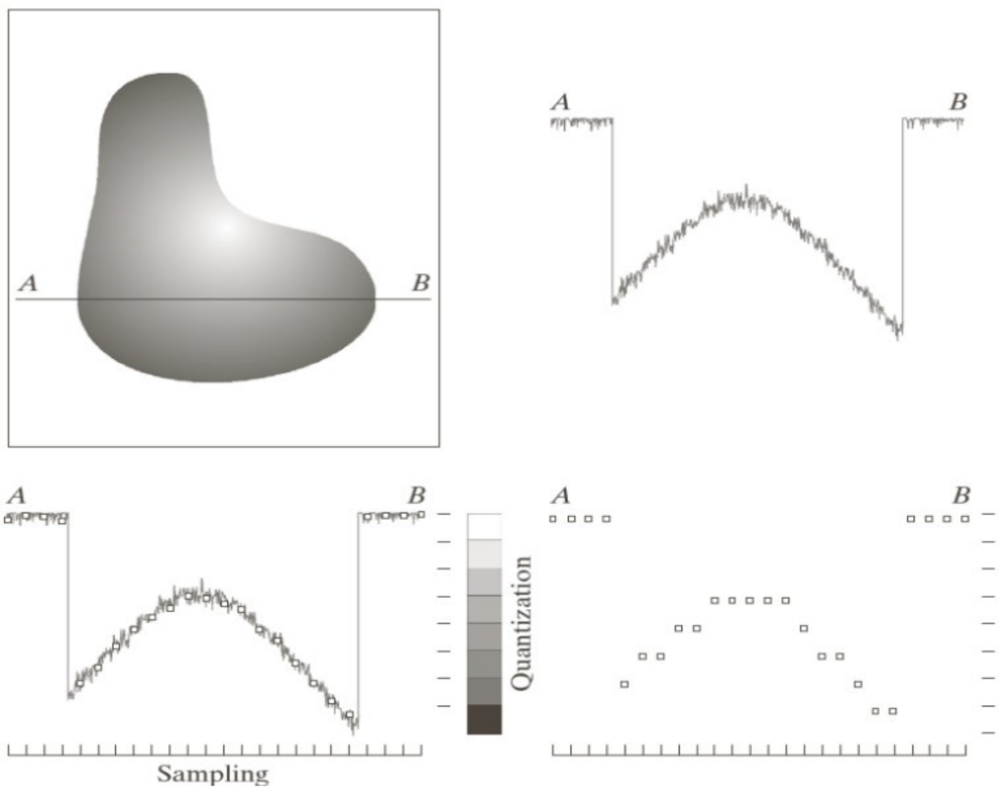


Image Transform

A particularly important class of 2-D linear transforms, denoted T(u, v)

$$T(u,v)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)r(x,y,u,v)$$

where $f(x,y)$ is the input image,

$r(x,y,u,v)$ is the *forward transformation kernel*,

variables u and v are the transform variables,

u = 0, 1, 2, ..., M-1 and v = 0, 1, ..., N-1.

Given T(u, v), the original image f(x, y) can be recovered using the inverse transformation of T(u, v).

$$f(x,y)=\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} T(u,v)s(x,y,u,v)$$

where $s(x,y,u,v)$ is the *inverse transformation kernel*,

x = 0, 1, 2, ..., M-1 and y = 0, 1, ..., N-1.

Forward Transform Kernel

$$T(u,v)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)r(x,y,u,v)$$

The kernel $r(x,y,u,v)$ is said to be SEPERABLE if
$r(x,y,u,v)=r_1(x,u)r_2(y,v)$

In addition, the kernel is said to be SYMMETRIC if
$r_1(x,u)$ is functionally equal to $r_2(y,v)$, so that
$r(x,y,u,v)=r_1(x,u)r_1(y,u)$

The Kernels for 2-D Fourier Transform

The *forward* kernel
$r(x,y,u,v)=e^{-j2\pi(ux/M+vy/N)}$
Where $j=\sqrt{-1}$

The *inverse* kernel
$$s(x,y,u,v)=\frac{1}{MN}e^{j2\pi(ux/M+vy/N)}$$

2-D Fourier Transform

$$T(u,v)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)e^{-j2\pi(ux/M+vy/N)}$$

$$f(x,y)=\frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}T(u,v)e^{j2\pi(ux/M+vy/N)}$$

NOISE MODELS, SEGMENTATION & COLOR IMAGE PROCESSING

Noise Models:

The sources of noise in digital images arise during image acquisition (digitization) and transmission. Imaging sensors can be affected by ambient conditions.Interference can be added to an image during transmission.

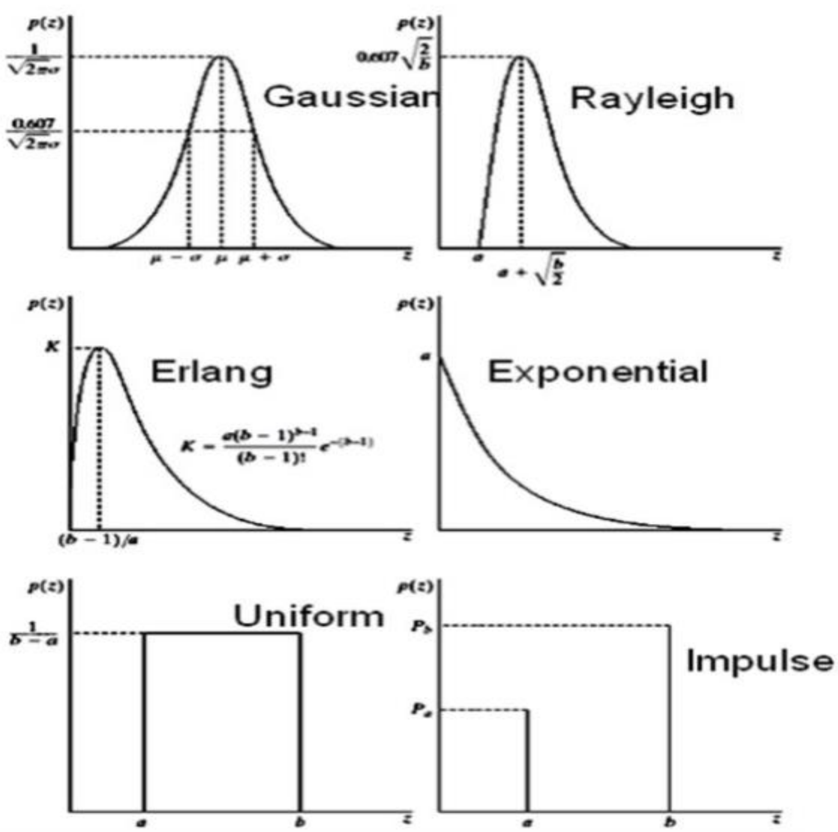
We can consider a noisy image to be modelled as follows:

$$g(x, y) = f(x, y) + \eta(x, y)$$

Where f(x, y) is the original image pixel, η(x, y) is the noise term and g(x, y) is the resulting noisy pixel.If we can estimate the noise model we can figure out how to restore the image.There are many different models for the image noise term η(x, y):

- Gaussian
  - Most common model
- Rayleigh
- Erlang (Gamma)
- Exponential
- Uniform
- Impulse

  Salt and pepper noise



Mean Filter:

The spatial filters of different kinds are used to remove different kinds of noise The arithmetic mean filter is a very simple one and is calculated as follows:

$$\hat{f}(x,y) = \frac{1}{mn} \sum_{(s,t)\in S_{xy}} g(s,t)$$

This is implemented as the simple smoothing filter.It blurs the image.
There are different kinds of mean filters all of which exhibit slightly different behaviour:

- Geometric Mean
- Harmonic Mean
- Contraharmonic Mean

Geometric Mean:

$$\hat{f}(x,y) = \left[\prod_{(s,t)\in S_{xy}} g(s,t) \right]^{\frac{1}{mn}}$$

Achieves similar smoothing to the arithmetic mean, but tends to lose less image detail.

Harmonic Mean:

$$\hat{f}(x,y) = \frac{mn}{\sum_{(s,t)\in S_{xy}} \frac{1}{g(s,t)}}$$

Works well for salt noise, but fails for pepper noise. Also does well for other kinds of noise such as Gaussian noise.

Contraharmonic Mean:

$$\hat{f}(x,y) = \frac{\sum_{(s,t)\in S_{xy}} g(s,t)^{Q+1}}{\sum_{(s,t)\in S_{xy}} g(s,t)^{Q}}$$

Q is the order of the filter. Positive values of Q eliminate pepper noise. Negative values of Q eliminate salt noise. It cannot eliminate both simultaneously.

Order Statics Filter:

Spatial filters based on ordering the pixel values that make up the neighbourhood defined by the filter support. Useful spatial filters include

- Median filter
- Max and min filter
- Midpoint filter
- Alpha trimmed mean filter

Adaptive filters:

The filters discussed so far are applied to an entire image without any regard for how image characteristics vary from one point to another.

- The behaviour of **adaptive filters** changes depending on the characteristics of the image inside the filter region.
- We will take a look at the **adaptive median filter.**
- The median filter performs relatively well on impulse noise as long as the spatial density of the impulse noise is not large.
- The adaptive median filter can handle much more spatially dense impulse noise, and also performs some smoothing for non-impulse noise.

The key to understanding the algorithm is to remember that the adaptive median filter has three purposes:

- Remove impulse noise
- Provide smoothing of other noise
- Reduce distortion (excessive thinning or thickenning of object boundaries).

In the adaptive median filter, the filter size changes depending on the characteristics of the image.

Notation:

- S_{xy} = the support of the filter centerd at *(x, y)*
- z_{min} = minimum grey level in S_{xy}
- z_{max} = maximum grey level in S_{xy}
- z_{med} = median of grey levels in S_{xy}
- z_{xy} = grey level at coordinates *(x, y)*

S_{max} =maximum allowed size of S_{xy}

Stage A: $A1 = z_{med} - z_{min}$

$A2 = z_{med} - z_{max}$

If $A1 > 0$ and $A2 < 0$, Go to stage B

Else increase the window size

If window size $\leq S_{max}$ repeat stage A

Else output z_{med}

Stage B: $B1 = z_{xy} - z_{min}$

$B2 = z_{xy} - z_{max}$

If $B1 > 0$ and $B2 < 0$, output z_{xy}

Else output z_{med}

Stage A determines if the output of the median filter z_{med} is an impulse or not (black or white). If it is not an impulse, we go to stage B.If it is an impulse the window size is increased until it reaches S_{max} or z_{med} is not an impulse.Note that there is no guarantee that z_{med} will not be an impulse. The smaller the the density of the noise is, and, the larger the support S_{max}, we expect not to have an impulse.

Stage B determines if the pixel value at *(x, y),* that is z_{xy}, is an impulse or not (black or white). If it is not an impulse, the algorithm outputs the unchanged pixel value z_{xy} .If it is an impulse the algorithm outputs the median z_{med} .

Band reject filters: Removing periodic noise form an image involves removing a particular range of frequencies from that image. Band reject filters can be used for this purpose.

An ideal band reject filter is given as follows:

$$H(u,v)=\begin{cases}1 & if\ D(u,v)<D_0-\frac{W}{2}\\ 0 & if\ D_0-\frac{W}{2}\le D(u,v)\le D_0+\frac{W}{2}\\ 1 & if\ D(u,v)>D_0+\frac{W}{2}\end{cases}$$

Band Pass filters:

Performs the opposite operation of band reject (BR) filters, which can be obtained from those of corresponding band reject filters

$$H_{BP}(u,v) = 1 - H_{BR}(u,v)$$

Notch Filters: Either rejecting (NR) or passing (NP) frequencies in predefined neighborhoods about a center frequency. The transfer functions of NP and NR have the following relationship.

$$H_{NP}(u,v) = 1 - H_{NR}(u,v)$$

Generally, interference components consist of multiple interference components. Traditional notch filters may remove too much information (unacceptable). The way-out of this problem is to use an optimum method by minimizing local variances of the restored estimate.

Optimum Notch Filter

When several interference components are present or if the interference has broad skirts a simply notch filter may remove too much image information. One solution is to use an optimum filter which minimizes local variances of the restored estimate

1. Manually place a notch pass filter HNP at each noise spike in the frequency domain. The Fourier transform of the interference noise pattern is

$$N(u,v) = H_{NP}(u,v)G(u,v$$

2. Determine the noise pattern in the spatial domain

$$\eta(x,y)=F^{-1}\{H(u,v)G(u,v)\}$$

3. Conventional thinking would be to simply eliminate noise by subtracting the periodic noise from the noisy image

4. To construct an optimal filter consider $\hat{f}(x,y) = g(x,y) - w(x,y)\,\eta(x,y)$

where w(x,y) is a weighting function.

5. We use the weighting function w(x,y) to minimize the variance of with respect to w(x,y)

$$w(x,y)=\frac{\overline{g(x,y)\eta(x,y)}-\overline{g}(x,y)\overline{\eta}(x,y)}{\overline{\eta^2}(x,y)-\overline{\eta}^2(x,y)}$$

We only need to compute this for one point in each non-overlapping neighborhood.

Inverse Filtering: To get F^(u,v) from degraded image G(u,v) :
-If degraded image is given by degradation + noise

$$G(u,v) = H(u,v)F(u,v) + N(u,v)$$

- Estimate the image by dividing by the degradation function H (u,v)

$$\hat{F}(u,v) = \frac{G(u,v)}{H(u,v)} = F(u,v) + \frac{N(u,v)}{H(u,v)}$$

We can never recover F(u,v) exactly:
1. N(u,v) is not known since □(x,y) is a random variable. — estimated
2. If H(u,v) ->0 then noise term will dominate. Helped by restricting analysis to (u,v) near origin.

Wiener (Minimum mean square Error) Filtering:

$$e^2 = E\left\{\left(f - \hat{f}\right)^2\right\} \rightarrow 0$$

$$\hat{F}(u,v) = \left[\frac{1}{H(u,v)} \times \frac{|H(u,v)|^2}{|H(u,v)|^2 + S_\eta(u,v)/S_f(u,v)}\right] G(u,v)$$

H(u,v) = degradation function
H*(u,v) = complex conjugate of H
|H(u,v)|= H*(u,v) H(u,v)
S(u,v)=$|N(u,v)|^2$=power spectrum of noise(estimated)
$S_f(u,v)=|F(u,v)|^2$=power spectrum of original image (not known)
If $S_f(u,v)$ is not known

$$e^2 = E\left\{\left(f - \hat{f}\right)^2\right\} \rightarrow 0$$

$$\hat{F}(u,v) = \left[\frac{1}{H(u,v)} \times \frac{|H(u,v)|^2}{|H(u,v)|^2 + K}\right] G(u,v)$$

Image Segmentation:

- Image segmentation divides an image into regions that are connected and have some similarity within the region and some difference between adjacent regions.
- The goal is usually to find individual objects in an image.
- For the most part there are fundamentally two kinds of approaches to segmentation: discontinuity and similarity.

  - Similarity may be due to pixel intensity, color or texture.
  - Differences are sudden changes (discontinuities) in any of these, but especially sudden changes in intensity along a boundary line, which is called an edge.
- There are three kinds of discontinuities of intensity: points, lines and edges.
- The most common way to look for discontinuities is to scan a small mask over the image. The mask determines which kind of discontinuity to look for.

$$R = w_1 z_1 + w_2 z_2 + \dots + w_9 z_9 = \sum_{i=1}^{9} w_i z_i$$

- Only slightly more common than point detection is to find a one pixel wide line in an image.

- For digital images the only three point straight lines are only horizontal, vertical, or diagonal (+ or –45°).

FIGURE 10.3 Line masks.

Horizontal			+45°			Vertical			–45°		
–1	–1	–1	–1	–1	2	–1	2	–1	2	–1	–1
2	2	2	–1	2	–1	–1	2	–1	–1	2	–1
–1	–1	–1	2	–1	–1	–1	2	–1	–1	–1	2

Detection of Discontinuities:

Gradient Operators:

First-order derivatives: The gradient of an image $f(x,y)$ at location (x,y) is defined as the vector:

$$\nabla \mathbf{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

The magnitude of this vector

$$\nabla f = \text{mag}(\nabla \mathbf{f}) = \left[G_x^2 + G_y^2\right]^{1/2}$$

The direction of this vector:

$$\alpha(x,y) = \tan^{-1}\left(\frac{G_x}{G_y}\right)$$

-Roberts cross-gradient operators
-Prewitt operators
-Sobel operators

−1	0
0	1

0	−1
1	0

Roberts

−1	−1	−1
0	0	0
1	1	1

−1	0	1
−1	0	1
−1	0	1

$$\nabla^2 h(r) = -\left[\frac{r^2-\sigma^2}{\sigma^4}\right] e^{-\frac{r^2}{2\sigma^2}}$$

Prewitt

−1	−2	−1
0	0	0
1	2	1

−1	0	1
−2	0	2
−1	0	1

Sobel

Second-order derivatives: (The Laplacian) The Laplacian of an 2D function f(x,y) is defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

- Consider the function:

$$h(r) = -e^{-\frac{r^2}{2\sigma^2}} \quad \text{where } r^2 = x^2 + y^2$$
$$\text{and } \sigma : \text{the standard deviation}$$

- The Laplacian of h is

$$\nabla^2 h(r) = -\left[\frac{r^2-\sigma^2}{\sigma^4}\right] e^{-\frac{r^2}{2\sigma^2}}$$

- The Laplacian of a Gaussian sometimes is called the Mexican hat function. It also can be computed by smoothing the image with the Gaussian smoothing mask, followed by application of the Laplacian mask.

Edge Linking and Boundary Detection

Two properties of edge points are useful for edge linking:
  – the strength (or magnitude) of the detected edge points
  – their directions (determined from gradient directions)
- This is usually done in local neighborhoods.
- Adjacent edge points with similar magnitude and direction are linked.
- For example, an edge pixel with coordinates (x_0,y_0) in a predefined neighborhood of (x,y) is similar to the pixel at (x,y) if

$$\left|\nabla f(x,y) - \nabla(x_0,y_0)\right| \le E, \quad E: \text{a nonnegative threshold}$$

$$\left|\alpha(x,y) - \alpha(x_0,y_0)\right| < A, \quad A: \text{a nonegative angle threshold}$$

Hough transform: a way of finding edge points in an image that lie along a straight line.
Example: *xy*-plane v.s. *ab*-plane (parameter space)

$$y_i = ax_i + b$$

- The Hough transform consists of finding all pairs of values of θ and ρ which satisfy the equations that pass through (*x*,*y*).
- These are accumulated in what is basically a 2-dimensional histogram.
- When plotted these pairs of θ and ρ will look like a sine wave. The process is repeated for all appropriate (*x*,*y*) locations.

Region-Based Segmentation

- Edges and thresholds sometimes do not give good results for segmentation.
- Region-based segmentation is based on the connectivity of similar pixels in a region.
  - Each region must be uniform.
  - Connectivity of the pixels within the region is very important.
- There are two main approaches to region-based segmentation: region growing and region splitting.

Basic Formulation

- Let *R* represent the entire image region.
- Segmentation is a process that partitions *R* into subregions, $R_1, R_2, \ldots, R_n$, such that

(a) $\bigcup R_i = R$

(b) R_i is a connected region, $i = 1,2,..., n$

(c) $R_i \cap R_j = \phi$ for all i and $j, i \neq j$

(d) $P(R_i) = \text{TRUE}$ for $i = 1,2,..., n$

(e) $P(R_i \cup R_j) = \text{FALSE}$ for any adjacent regions R_i and R_j

where $P(R_k)$: a logical predicate defined over the points in set R_k
For example: $P(R_k)$=TRUE if all pixels in R_k have the same gray level.

Region splitting is the opposite of region growing.

- First there is a large region (possible the entire image).
- Then a predicate (measurement) is used to determine if the region is uniform.
- If not, then the method requires that the region be split into two regions.
- Then each of these two regions is independently tested by the predicate (measurement).
- This procedure continues until all resulting regions are uniform.

The main problem with region splitting is determining where to split a region. One method to divide a region is to use a quad tree structure.Quadtree: a tree in which nodes have exactly four descendants. The split and merge procedure:

- Split into four disjoint quadrants any region R_i for which $P(R_i)$ = FALSE.
- Merge any adjacent regions R_j and R_k for which $P(R_j U R_k)$ = TRUE. (the quadtree structure may not be preserved)
- Stop when no further merging or splitting is possible.

Morphological image processing

- **Morphology**: a branch of biology that deals with the form and structure of animals and plants

- Morphological image processing is used to extract image components for representation and description of region shape, such as boundaries, skeletons, and the convex hull

Reflection

The reflection of a set B, denoted B, is defined as

$$B = \{w \mid w = -b, \text{for } b \in B\}$$

Translation

The translation of a set B by point $z = (z_1, z_2)$, denoted $(B)_Z$, is defined as

$$(B)_Z = \{c \mid c = b + z, \text{for } b \in B\}$$

Erosion

With A and B as sets in Z^2, the erosion of A by B, denoted $A \quad B$, defined

$$A \quad B = \{z \mid (B)_Z \subseteq A\}$$

The set of all points z such that B, translated by z, is contained by A.

$$A \quad B = \{z \mid (B)_Z \cap A^c = \varnothing\}$$

Dilation

With A and B as sets in Z^2, the dilation of A by B, denoted $A \oplus B$, is defined as

$$\mathrm{A} \oplus \mathrm{B} = \left\{z \mid \left(B\right)_z \cap A \neq \varnothing\right\}$$

The set of all displacements z, the translated B and A overlap by at least one element.

$$A \oplus B = \left\{z \mid \left[\left(B\right)_z \cap A\right] \subseteq A\right\}$$

Colour image processing

Colour fundamentals: In 1666 Sir Isaac Newton discovered that when a beam of sunlight passes through a glass prism, the emerging beam is split into a spectrum of colours. Chromatic light spans the electromagnetic spectrum from approximately 400 to 700 nm

3 basic qualities are used to describe the quality of a chromatic light source:

- **Radiance:** the total amount of energy that flows from the light source (measured in watts)
- **Luminance:** the amount of energy an observer *perceives* from the light source (measured in lumens).Note we can have high radiance, but low luminance
- **Brightness:** a subjective (practically unmeasurable) notion that embodies the intensity of light.

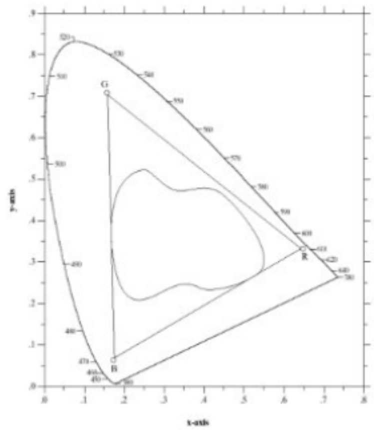
CIE chromacity diagram

Specifying colours systematically can be achieved using the CIE **chromacity diagram**. On this diagram the x-axis represents the proportion of red and the y-axis represents the proportion of red used.

The proportion of blue used in a colour is calculated as:

$$z = 1 - (x + y)$$

Any colour located on the boundary of the chromacity chart is fully saturated. The point of equal energy has equal amounts of each colour and is the CIE standard for pure white. Any straight line joining two points in the diagram defines all of the different colours that can be obtained by combining these two colours additively. This can be easily extended to three points.



This means the entire colour range cannot be displayed based on any three colours

The triangle shows the typical colour gamut produced by RGB monitors

The strange shape is the gamut achieved by high quality colour printers

Colour models:

Colour models provide a standard way to specify a particular colour, by defining a 3D coordinate system, and a subspace that contains all constructible colours within a particular model. Any colour that can be specified using a model will correspond to a single point within the subspace it defines. Each colour model is oriented towards either specific hardware (RGB, CMY, YIQ), or image processing applications (HSI).

The RGB Model

In the RGB model each colour appears in its primary spectral components of red, green and blue. The model is based on a Cartesian coordinate system

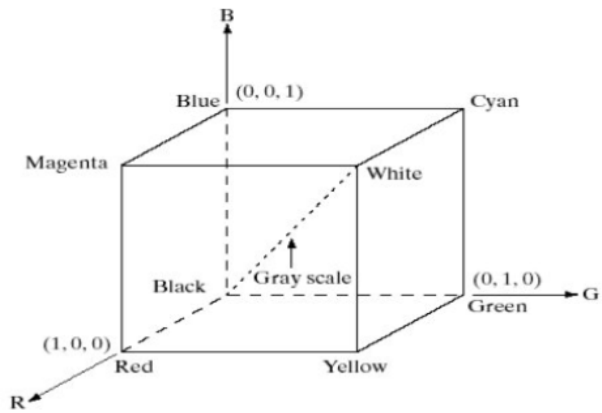
RGB values are at 3 corners

Cyan magenta and yellow are at three other corners

Black is at the origin

White is the corner furthest from the origin

Different colours are points on or inside the cube represented by RGB vectors



Images represented in the RGB colour model consist of three component images – one for each primary colour.When fed into a monitor these images are combined to create a composite colour image.The number of bits used to represent each pixel is referred to as the colour depth.A 24-bit image is often referred to as a full-colour image as it allows16,777,216 colours .The RGB model is used for colour monitors and most video cameras.

The CMY Model

The CMY (cyan-magenta-yellow) model is a *subtractive* model appropriate to absorption of colours, for example due to pigments in paints. Whereas the RGB model asks what is added to black to get a particular colour, the CMY model asks what is subtracted from white. In this case, the primaries are cyan, magenta and yellow, with red, green and blue as secondary colours .When a surface coated with cyan pigment is illuminated by white light, no red light is reflected, and similarly for magenta and green, and yellow and blue. The CMY model is used by printing devices and filters. Equal amounts of the pigment primaries, cyan, magenta, and yellow should produce black. In practice, combining these colors for printing produces a muddy-looking black. To produce true black, the predominant color in printing, the fourth color, black, is added, giving rise to the CMYK color model.

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The HSI Model

The HSI model uses three measures to describe colours:

Hue: A colour attribute that describes a pure colour (pure yellow, orange or red)

Saturation: Gives a measure of how much a pure colour is diluted with white light

Intensity: Brightness is nearly impossible to measure because it is so subjective. Instead we use intensity. Intensity is the same achromatic notion that we have seen in grey level images.

Colour Image Quantization

Color quantization or color image quantization is a process that reduces the number of distinct colors used in an image, usually with the intention that the new image should be as visually similar as possible to the original image. Computer algorithms to perform color quantization on bitmaps have been studied since the 1970s. Color quantization is critical for displaying images with many colors on devices that can only display a limited number of colors, usually due to memory limitations, and enables efficient compression of certain types of images. The name "color quantization" is primarily used in computer graphics research literature; in applications, terms such as optimized palette generation, optimal palette generation, or decreasing color depth are used.

Histogram of Colour Image

A histogram is a graphical representation of the number of pixels in an image. In a more simple way to explain, a histogram is a bar graph, whose X-axis represents the tonal scale(black at the left and white at the right), and Y-axis represents the number of pixels in an image in a certain area of the tonal scale. For example, the graph of a luminance histogram shows the number of pixels for each brightness level(from black to white), and when there are more pixels, the peak at the certain luminance level is higher.

A color histogram of an image represents the distribution of the composition of colors in the image. It shows different types of colors appeared and the number of pixels in each type of the colors appeared. The relation between a color histogram and a luminance histogram is that a color histogram can be also expressed as “Three Color Histograms”, each of which shows the brightness distribution of each individual Red/Green/Blue color channel.

Principles of the formation of a color histogram

The formation of a color histogram is rather simple. From the definition above, we can simply count the number of pixels for each 256 scales in each of the 3 RGB channel, and plot them on 3 individual bar graphs.In general, a color histogram is based on a certain color space, such as RGB or HSV. When we compute the pixels of different colors in an image, if the color space is large, then we can first divide the color space into certain numbers of small intervals. Each of the intervals is called a bin. This process is called color quantization. Then, by counting the number of pixels in each of the bins, we get the color histogram of the image.

Colour Image Segmentation

Segmentation in RGB Vector Space

Let the average color of interest is denoted by the RGB vector a. Let z denote an arbitrary point in RGB space.

$$D(z,a) = \|z-a\| = \left[(z-a)^T(z-a)\right]^{1/2}$$
$$= \left[(z_R-a_R)^2 + (z_G-a_G)^2 + (z_B-a_B)^2\right]^{1/2}$$