GCA S6 01
Exam Code: IP

Image Processing

SEMESTER-VI

BACHELOR OF COMPUTER APPLICATION

BLOCK- 2



KRISHNA KANTA HANDIQUI STATE OPEN UNIVERSITY

Subject Experts

1. **Prof. Anjana Kakati Mahanta**, Gauhati University
2. **Prof. (Retd.) Pranhari Talukdar**, Gauhati University
3. **Dr. Jyotiprokash Goswami**, Associate Professor, Assam Engineering College

Course Co-ordinator : Dr. Tapashi Kashyap Das, Assistant Professor, KKHSOU
Ms. Sruti Sruba Bharali, Assistant Professor, KKHSOU

SLM Preparation Team

UNITS	CONTRIBUTORS
8	**Dr.Chandan Kumar Nath**, Jr.Tech. Superintendent, IITG
9,10 & 11	**Sangeeta Borkakoty**, USTM
12 & 14	**Daisy Kalita,** USTM
13 & 15	**Dr. Minakshi Gogoi,** USTM

Editorial Team

Content	:	**Dr.Sanjib Kr. Kalita**, Gauhati University
Language	:	**Prof. (Retd.) Robin Goswami**, Cotton College
Structure, Format & Graphics	:	**Dr.Tapashi Kashyap Das**, KKHSOU

December, 2019 **ISBN:**

Copyright © Krishna KantaHandiqui State Open University

This Self Learning Material (SLM) of the Krishna Kanta Handiqui State University is made available under a Creative Commons Attribution-Non Commercial-ShareAlike4.0 License (International): http.//creativecommons.org/licenses/by-nc-sa/4.0.

Printed and published by Registrar on behalf of the Krishna Kanta Handiqui State Open University.

Headquarter : Patgaon, Rani Gate, Guwahati - 781017; Web : www.kkhsou.in/web_new
City Office : Housefed Complex, Dispur, Guwahati-781006;

The University acknowledges with thanks the financial support provided by the **Distance Education Bureau**, **UGC** for the preparation of this study material.

CONTENTS

		Pages
UNIT 8:	**Image Enhancement in Frequency Domain** Introduction, Low pass filtering in frequency domain, High pass filter in frequency domain, Butterworth high pass filter, Homomorphic filter.	**5–9**
UNIT 9:	**Image Restoration** Introduction, Image degradation, Types of image Blur, Classification of Image Restoration technique, Image restoration model, Linear Image Restoration techniques, Pseudo Inverse filter, Wiener filter, Non Linear Image Restoration techniques : Iterative method, Maximum likelihood Method, Stochastic image restoration technique.	**10–28**
UNIT 10:	**Image Noise** Introduction, classification of noise image: adaptive noise, Multiplicative noise, impulse noise, Median filtering, trimmed average filter.	**29–37**
UNIT 11:	**Image Segmentation** Introduction, Classification of Image segmentation technique, Local and global segmentation, region approach to image segmentation, clustering technique, Image segmentation based on thresholding, Limitation, edge detection techniques, edge linking, Watershed transformation.	**38–51**
UNIT 12:	**Image Compression** Introduction, Need for Image compression, Redundancy in images, Image compression scheme Huffman coding, Arithmetic coding, transform base coding, Vector quantization.	**52–68**
UNIT 13:	**Binary Image Processing** Introduction, Binarisation, Mathematical morphology, structuring element, Morphological image processing, Basic set theory, Logical operations, Standard binary morphological operations : Dialation, Erosion, Opening, Closing, Properties of Morphological operations.	**69–82**
UNIT 14:	**Color Image Processing** Introduction, Color formation, Human perception of color, Color model: RGB model, CMY color model, HIS color model, YIQ color model, Histogram equalization of color image, Color Image filtering.	**83–102**
UNIT 15:	**Wavelet-based Image Processing** Introduction, Wavelet, Wavelet transform, 2D continuous wavelet transform, Discrete wavelet transform, Haar Wavelet, Filter bank, Embedded Image coding.	**103–111**

BLOCK INTRODUCTION

This is the second block of the course ***Image Processing.*** This block comprises the following eight units :

Unit 8 provides an overview of image enhancement in frequency domain. Frequency refers to the rate of repetition of some periodic event. It is possible to perform filtering in frequency domain by specifying the frequencies that should be kept and the frequencies that should be discarded. The process of selective frequency inclusion or exclusion is termed frequency domain filtering. **Unit 9** describe the image restoration. Image Restoration is the operation of taking a corrupt/noisy image and estimating the clean, original image. Corruption may come in many forms suh as motion blur, noise and camera mis-focus. **Unit 10** describes the image noise. Noise means unwanted information in the image content. Different noise distribution statistical models have been discussed in this unit. Image segmentation has been explained in **Unit 11**. It refers to the process of partitioning an image into groups of pixels which are homogeneous with respect to some criterion. Different groups must not intersect each other and adjacent groups must be heterogeneous. Segmentation algorithms are area oriented instead of pixel oriented. **Unit 12** has been described the concept of image compression techniques, needs and applications. Different compression techniques like Huffman coding, Arithmetic coding etc. have been explained with examples. **Unit 13** is concerned with binary image processing. The basic type of image is binary image, which requires only 1 bit per pixel. Set theory operations like union, intersection etc. are implemented in binary image and have been explained in this unit. **Unit 14** describes the colour image processing. Now-a-days colour images occupy more space when compared to grayscale or binary images. Different colour models have been explained in this unit. **Unit 15** explain the wavelet based image processing. The disadvantages of Fourier transforms has been discussed in this unit and how it is overcome using wavelet based image transform also presented.

Each unit of this block includes some along-side boxes to help you know some of the difficult, unseen terms. Some “EXERCISES” have been included to help you apply your own thoughts. You may find some boxes marked with: “LET US KNOW”. These boxes will provide you with some additional interesting and relevant information. Again, you will get “CHECK YOUR PROGRESS” questions. These have been designed for you to self-check your progress of study. It will be helpful for you if you solve the problems put in these boxes immediately after you go through the sections of the units and then match your answers with “ANSWERS TO CHECK YOUR PROGRESS” given at the end of each unit.

UNIT 8 : IMAGE ENHANCEMENT IN FREQUENCY DOMAIN

UNIT STRUCTURE

8.1 Learning Objectives

8.2 Introduction

8.3 Low Pass Filtering

8.4 High Pass Filtering

8.5 Butterworth High Pass Filter

8.6 Homomorphic Filter

8.7 Let Us Sum Up

8.8 Further Reading

8.9 Answers to Check Your Progress

8.10 Model Questions

8.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- explain different image filters
- learn about low and high pass image filter
- list the drawbacks of low and high filter
- define butterworth high pass filter
- describe homomorphic filter and its use

8.2 INTRODUCTION

Image enhancement is the process of digitally manipulating a stored image using different techniques for better visualization or interpretation of the image. Image enhancement not only provides means for improved perception but also helps in the further processing of the digital image. We have studied image enhancement in spatial domain in Unit 5. In spatial domain, certain operations are carried out on the pixels for achieving the desired output image, but in frequency domain Fourier transform of the image is taken and multiply the result by a filter, and

take the inverse transform to produce the enhanced image. The idea is to increase or decrease the high or low frequency components of an image and getting the blurred or sharpened image as a result.

8.3 LOW PASS FILTERING

In low pass filtering method high frequency component of an image are eliminated and it results in blurring of the image and, thus, a reducing the sharp transitions of noise. A low frequency is set as a cut-off and the Fourier transform or an image is let to pass through the cut-off. The frequency components higher than the cut-off cannot pass through the filter and as a result we get an image with low frequency level. An ideal low pass filter can be defined as follows:

$$H(u,v) = \begin{cases} 1, \text{ if } \sqrt{u^2+v^2} \le r_0 \\ 0, \text{ if } \sqrt{u^2+v^2} > r_0 \end{cases}$$

where r_0 is the cut-off frequency.

$\sqrt{u^2+v^2}$ is the distance of a point *(u, v)* from the origin of the frequency plane. The important factor determining the quality of low pass filtering result is the choice of cut-off suitable to the intended application of the images. In practice, it is almost impossible to eliminate all the high frequency components and two problems arise when performing a low pass filtering: blurring and ringing.

8.4 HIGH PASS FILTERING

In short, we can say that high pass filter is just the opposite of low pass filter. In high pass filtering technique low frequency components are eliminated and a sharpened image is obtained. A high frequency is set as cut-off and the frequency components which are lower than this cut-off are eliminated and thus resulting in a sharpened image. An ideal high pass filter is the one which eliminates almost all the low frequency components and it is defined as

$$H(u,v) = \begin{cases} 0, \text{ if } \sqrt{u^2+v^2} \le r_0 \\ 1, \text{ if } \sqrt{u^2+v^2} > r_0 \end{cases}$$

One of the major drawbacks of high pass filtering is that the noise present in an image may get amplified. Ringing effect is also present in image passed through high pass filter. Butterworth high pass filter tries to minimize these problems in a novel way.



CHECK YOUR PROGRESS

Q 1: Mention two problems caused in low pass filter?

.....

.....

Q 2: Differentiate between low pass filter and high pass filter?

.....

.....

Q 3: Write one application of high pas filter in image processing methods?

.....

.....

8.5 BUTTERWORTH HIGH PASS FILTER

Butterworth high-pass filter controls the sharpness of the filter based on its order. The Butterworth function is defined as:

$$H(u,v) = \frac{1}{1+\left(\frac{D_0}{\sqrt{u^2+v^2}}\right)^{2n}}$$

where, D_0 is the cut-off distance between a point *(u, v)* and the origin of frequency plane and n is the order of the filter. Butterworth high pass filter has a gradual transition from 0 to 1 and thus reduces ringing effect.

8.6 HOMOMORPHIC FILTER

Homomorphic filter is another filter which is used in image enhancement. Homomorphic filtering method involves a nonlinear mapping to a different domain in which linear filter techniques are applied, followed

by mapping back to the original domain. In case of image processing, the light intensity function involving multiplication of reflectance and illumination cannot be represented in Fourier series. Instead, logarithmic values of the each function are taken and the light intensity function is represented as the sum of the logarithmic values. A high pass filter is applied to the Fourier transform of the image. Homomorphic is defined as

$$H(u,v) = (\gamma_H - \gamma_L)\left\{1 - \exp\left\{-c\left\{\frac{D(u_2 v)}{D_0}\right\}^2\right\}\right\} + \gamma_L$$

$\tilde{a}_H$ is the high frequency gain and $\tilde{a}_L$ is the low frequency gain. *c* is a constant and typically its value is 0.5. Homomorphic filter usually increases the contrast of an image.



CHECK YOUR PROGRESS

Q 4: Which effect is reduced by applying Butterworth filter?

.....

.....

Q 5: Why logarithmic values are taken in the Homomorphic filtering method?

.....

.....



8.7 LET US SUM UP

- Image filters are used for enhancement of image.
- Low pass filter helps in reducing the noise present in an image.
- High pass filter enhances the edges the of object in an image and helps in segmenting them.
- Butterworth filter enhances image quality without causing ringing effect.
- Homomorphic filter usually enhances the contrast of an image.



8.8 FURTHER READING

1. Gonzales, R. C., & Woods, R. E. (2008). Digital image processing. Prentice Hall.
2. Jayaraman, S. Esakkirajandan, S., & Veerakumar, T. (2009). Digital Image Processing, New Delhi: Tata McGraw-Hill Education Private Limited.

8.9 ANSWERS TO CHECK YOUR PROGRESS

Ans to Q No 1: Blurring and ringing effect.

Ans to Q No 2: Low pass filter eliminates the high frequency components of an image, whereas high pass filter does the opposite. Low pass filter reduces noise but high pass filter amplify the noise.

Ans to Q No 3: Edge detection.

Ans to Q No 4: Ringing effect.

Ans to Q No 5: Fourier transform cannot be represented as the product of two separate function. But light intensity function is a product of two functions and thus we have to take the logarithmic values and represent them as sum of functions.



8.10 MODEL QUESTIONS

Q 1: Describe low pass filter method and its effect on the resulting image.

Q 2: What is an ideal high pass filter?

Q 3: Why Butterworth high pass filter is efficient compared to high pass filter?

Q 4: What property of an image is enhanced by Homomorphic filter?

Q 5: Write the equation for homomorphic filter and describe the variables. What is the typical value of the constant in Homomorphic filter equation?

*** ***** ***

UNIT 9 : IMAGE RESTORATION

UNIT STRUCTURE

9.1 Learning Objectives

9.2 Introduction

9.3 Image Degradation

9.4 Types of Image Blur

9.5 Image Restoration Model

9.6 Noise

- 9.6.1 Sources of Noise
- 9.6.1 Noise Models

9.7 Classification of Image Restoration Techniques

9.8 Deterministic Image Restoration Techniques

- 9.8.1 Linear Image Restoration Techniques
- 9.8.2 Inverse Filter and Pseudo Inverse Filter
- 9.8.3 Wiener Filter
- 9.8.4 Non Linear Image Restoration Techniques
- 9.8.5 Iterative Method
- 9.8.6 Maximum Likelihood Method

9.9 Stochastic image restoration technique

9.10 Let Us Sum Up

9.11 Further Reading

9.12 Answers to Check Your Progress

9.13 Model Questions

9.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- learn about the image degradation
- define noise, its sources and types
- describe the various image restoration techniques

9.2 INTRODUCTION

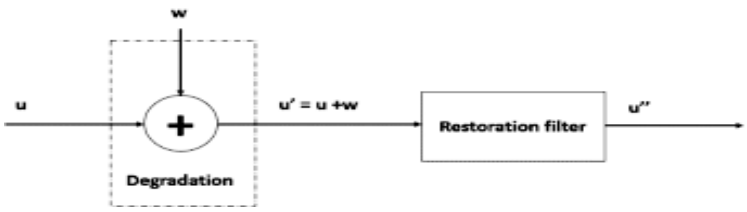
Image restoration is the operation of taking a corrupt/noisy image and estimating the clean, original image. Corruption may come in many forms such as motion blur, noise and camera mis-focus. Image restoration is performed by reversing the process that blurred the image and is performed by imaging a point source and using the point source image, which is called the Point Spread Function (PSF) to restore the image information lost to the blurring process.

Image restoration is different from image enhancement in that image enhancement is largely a subjective process, while image restoration is for objective process. Restoration attempts to recover an image that has been degraded by using a prior knowledge of the degradation phenomenon. Thus, restoration techniques are oriented modeling the degradation and applying the inverse process in order to recover the original image. The objective of image restoration techniques is to reduce noise and to recover the resolution loss. There are different types of image restoration techniques, each with their own advantages and demerits.

9.3 IMAGE DEGRADATION

In practice, an image may be degraded by various types and forms of noise. However, the most common type of noise is the additive one. Figure 9.1 shows, the degradation process modelled as an additive noise term where w operates on an input image u, to produce a degraded image u''. Given this noisy observation, along with some knowledge of the additive noise term, the restoration technique yields an estimate u'', of the original image. The denoised estimate is desired to be as close as possible to original image u.

Figure 9.1: Image Degradation



9.4 TYPES OF IMAGE BLUR

Blur is happening when the edge brightness of an image is reduced and the transition from one color to the other is very smooth. There are different types of image blur:

- **Defocus aberration**, blurring of an image due to incorrect focus

  In optics, defocus is the aberration in which an image is simply out of focus. This aberration is familiar to anyone who has used a camera, videocamera, microscope, telescope, or binoculars. Optically, defocus refers to a translation of the focus along the optical axis away from the detection surface. In general, defocus reduces the sharpness and contrast of the image. What should be sharp, high-contrast edges in a scene become gradual transitions. Fine detail in the scene is blurred or even becomes invisible. Nearly all image-forming optical devices incorporate some form of focus adjustment to minimize defocus and maximize the image quality.

- **Motion blur**, blurring of an image due to movement of the subject or imaging system. Motion blur is the apparent streaking of moving objects in a photograph or a sequence of frames, such as a film or animation. It results when the image being recorded changes during the recording of a single exposure due to rapid movement or long exposure.

- **Bokeh**, the aesthetic quality of the out-of-focus parts of an image. In photography, bokeh is the aesthetic quality of the blur produced in the out-of-focus parts of an image produced by a lens. Bokeh has been defined as “the way the lens renders out-of-focus points of light”. Differences in lens aberrations and aperture shape cause some lens designs to blur the image in a way that is pleasing to the eye, while others produce blurring that is unpleasant or distracting (“good” and “bad” bokeh, respectively). Bokeh occurs for parts of the scene that lie outside the depth of field. Photographers sometimes deliberately use a shallow focus technique to create images with prominent out-of-focus regions.

- **Box blur**, a graphic-art effect. A box blur (also known as a box linear filter) is a spatial domain linear filter in which each pixel in the resulting image has a value equal to the average value of its neighboring pixels in the input image. It is a form of low-pass ("blurring") filter.
- **Gaussian blur**, a graphic-art effect. In image processing, a Gaussian blur (also known as Gaussian smoothing) is the result of blurring an image by a Gaussian function. It is a widely used effect in graphics software, typically to reduce image noise and to reduce detail. The visual effect of this blurring technique is a smooth blur resembling that of viewing the image through a translucent screen, distinctly different from the bokeh effect produced by an out-of-focus lens or the shadow of an object under usual illumination.



CHECK YOUR PROGRESS

Q 1: Image Restoration and Image Enhancement are the same. True/False?

Q 2: What is the main objective of Image Restoration?

.....

.....

Q 3: What is the most common type of noise?

.....

.....

Q 4: What is blurring of an image due to incorrect focus called?

.....

.....

Q 5: Why is Gaussian blur generally used?

.....

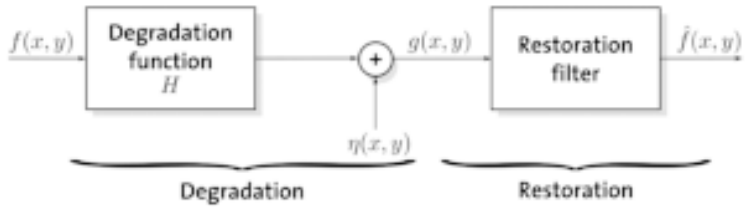
.....

9.5 IMAGE RESTORATION MODEL

Image restoration is the process of recovering an image that has been degraded by some knowledge of degradation function **H** and the

additive noise term $\eta(x,y)$. Thus in restoration, degradation is modelled and its inverse process is applied to recover the original image. Fig.9.2 shows image degradation and restoration model.

Figure 9.2: Model of image degradation/restoration



The objective of image restoration is to obtain an estimate of the original image *f(x,y)*. Here, by some knowledge of **H** and *(x,y)*, we find the appropriate restoration filters, so that output image *f'(x,y)* is as close to the original image *f(x,y)* as possible since it is practically not possible (or very difficult) to completely restore the original image.

Here,

g(x,y) - degraded image

f(x,y) - input or original image

f'(x,y) - recovered or restored image

(x,y) - additive noise term

9.6 NOISE

Noise is normally a high frequency component and a low pass filtering technique are used to remove the noise. Noise is introduced in the image at the time of image acquisition or transmission. There are different factors available for introducing noise in the image. The quantification of the noise will be decided by number of pixels corrupted in the image.

9.6.1 Sources of Noise

The principal sources of noise in the digital image are:

- The imaging sensor may be affected by environmental conditions during image acquisition.
- Insufficient light levels and sensor temperature may introduce the noise in the image.

- Interference in the transmission channel may also corrupt the image.
- If dust particles are present on the scanner screen, they can also introduce noise in the image.
- Blur due to miss-focus.
- Blur due to motion.
- Slow shutter speeds.
- High sensitivity modes (High ISO setting).
- Low light (night photos or dark scenes)

9.6.2 Noise Models

Noise is the undesirable elements produced in the image. During image acquisition or transmission, several factors are responsible for introducing noise in the image. Depending on the type of disturbance, the noise can affect the image to different extent. Generally, researchers keep focus on removing certain kind of noise. Therefore, researchers identify certain kind of noise and apply different algorithms to eliminate the noise. Image noise can be categorised as Impulse noise (Salt-and-pepper noise), Amplifier noise (Gaussian noise), Shot noise, Quantization noise (uniform noise), Film grain, on-isotropic noise, Multiplicative noise (Speckle noise) and Periodic noise.

Noise tells unwanted information in digital images. Noise produces undesirable effects such as artefacts, unrealistic edges, unseen lines, corners, blurred objects and disturbs background scenes. To diminish these undesirable effects, prior knowledge of noise models is required for further processing. Digital noise may arise from various kinds of sources such as Charge Coupled Device (CCD) and Complementary Metal Oxide Semiconductor (CMOS) sensors. In some sense, points spreading function (PSF) and modulation transfer function (MTF) have been used for timely, complete and quantitative analysis of noise models. Probability density unction (PDF) or Histogram is also used to design and

characterize the noise models.

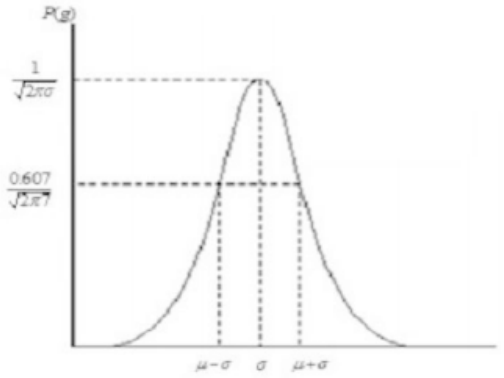
To understand the restoration process, it is imperative to understand the noise function (x, y). A good restoration system can be built if researchers could have realistic noise models. Noise in imaging systems are usually either additive or multiplicative. Noise is described by its probability characteristics. Some noise models are:

- **Gaussian Noise Model**

  Gaussian noise is caused by natural sources such as thermal vibration of atoms and discrete nature of radiation of warm objects. Gaussian noise generally disturbs the gray values in digital images. That is why Gaussian noise model essentially designed and characterised by its PDF or normalizes histogram with respect to gray value. This is given as:

  where g = gray value, ó = standard deviation and μ = mean. Generally, Gaussian noise mathematical model represents the correct approximation of real world scenarios. In this noise model, the mean value is zero; variance is 0.1 and 256 gray levels in terms of its PDF, which is shown in Fig.9.3.

Figure 9.3: PDF of Gaussian Noise



Due to this equal randomness the normalized Gaussian noise curve look like in bell shaped. The PDF of this noise model shows that 70% to 90% noisy pixel values of degraded

image in between and $\mu\sigma\mu\sigma -+$. The shape of normalized histogram is almost same in spectral domain.

- **White Noise**

Noise is essentially identified by the noise power. Noise power spectrum is constant in white noise. This noise power is equivalent to power spectral density function. The range of total noise power is to available in white noise in frequency domain. That means, ideally noise power is infinite in white noise. This fact is fully true because the light emitted from the sun has all the frequency components.

In white noise, correlation is not possible because every pixel values are different from their neighbors. That is why autocorrelation is zero. So that image pixel values are normally disturbed positively due to white noise.

- **Rayleigh Noise**

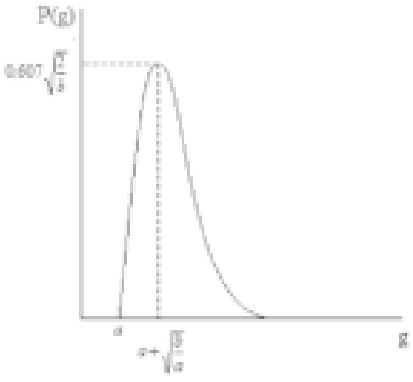
Unlike the Gaussian distribution, the Rayleigh distribution is not symmetric. Rayleigh noise is present in radar range images. In Rayleigh noise, probability density function is given as:

Figure 9.4: Rayleigh distribution

$$P(g) = \begin{cases} \frac{2}{b}(g-a)e^{\frac{-(g-a)^2}{b}} & for\ g \geq a \\ 0 & for\ g < a \end{cases}$$

Where mean $\mu = a + \sqrt{\frac{\pi b}{4}}$ and variance

$\sigma^2 = \frac{b(4-\pi)}{4}$ are given as, respectively.



Apart from these there are some other noise models as well, among which the prominent ones are:

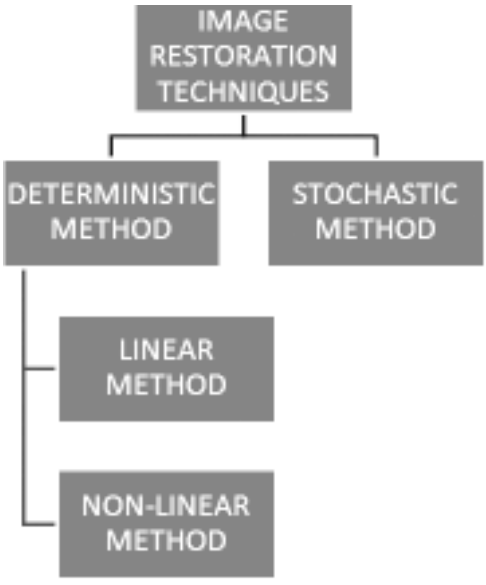
- **Brownian Noise:** Colored noise has many names such as Brownian noise or pink noise or flicker noise or 1/f noise. In Brownian noise, power spectral density is proportional to square of frequency over an octave i.e., its power falls on ¼th part (6 dB per octave). Brownian noise is caused by Brownian motion. Brownian motion is seen due to the random movement of suspended particles in fluid. Brownian noise can also be generated from white noise.
- **Impulse Valued Noise (Salt and Pepper Noise):** This noise appears as black and white dots on the entire image. This is also called data drop noise because statistically it drops the original data values. This noise is also referred to as salt and pepper noise. However, the image is not fully corrupted by salt and pepper noise instead some pixel values are changed in the image.
- **Periodic Noise (Sinusoidal Noise):** It is a special kind of noise that cannot be eliminated in the spatial domain. Images are usually corrupted by the periodic noise due to electrical and electro mechanical interface during the acquisition. It is virtually impossible to remove periodic noise using spatial domain masks. The only way to eliminate them is by studying their response in the frequency domain
- **Speckle Noise:** This noise is multiplicative noise. Their appearance is seen in coherent imaging system such as laser, radar and acoustics etc.
- **Quantization noise:** Quantization noise appearance is inherent in amplitude quantization process. It is generally present due to analog data converted into digital data.

9.7 CLASSIFICATION OF IMAGE RESTORATION TECHNIQUES

Image restoration techniques are the methods which effort the inversion of several degrading process. Image restoration technique can be broadly classified into two types depending upon the knowledge of degradation. If the previous knowledge concerning degradation is known subsequently, the deterministic method of image restoration can be employed. The blur may be due to a number of reasons, such as motion defocusing, and atmospheric turbulence the noise may originate in. Restoration techniques model the degradation process and attempt to apply an inverse procedure to obtain an approximation of the original image.

Many image restoration algorithms have their roots in well developed areas of mathematics such as estimation theory, the solution of ill posed problems, linear algebra and numerical analysis. Iterative image restoration techniques often attempt to restore an image linearly or non-linearly by minimizing some measures of degradation such as maximum likelihood, constrained least square, etc. Blind restoration techniques attempt to solve the restoration problem without knowing the blurring function.

Figure 9.5: Image Restoration Techniques



9.8 DETERMINISTIC IMAGE RESTORATION TECHNIQUES

If the prior knowledge about degradation is known then the deterministic method of image restoration can be employed.Deterministic techniques are of two types:

- linear and
- non-linear

9.8.1 Linear Image Restoration Techniques

If researchers know the exact PSF (Point Spread Function) model in the image degradation system, then the noise effect can be easily ignored. The main drawback of this filter is, it will not perform well in the presence of noise. It will tend to amplify the noise.

The most common, simplest and fastest kind of filtering is achieved by linear filters. The linear filter replaces each pixel with a linear combination of its neighbors and convolution kernel is used in prescription for the linear combination.

of the input signal x(n) with the impulse response h(n) of the given filter, i.e. the filter output arising from the input of an ideal Dirac impulse . When researchers apply linear filtering then each pixel is replaced by linear combination of its neighbor.

9.8.2 Inverse Filter and Pseudo Inverse Filter

Inverse Filter helps to get rid of the blurring (Deconvolution). The inverse filtering is a restoration technique for deconvolution, i.e., when the image is blurred by a known low pass filter, it is possible to recover the image by inverse filtering or generalized inverse filtering. However, inverse filtering is very

sensitive to additive noise. The approach of reducing one degradation at a time allows us to develop a restoration algorithm for each type of degradation and simply to combine them.

The Wiener filtering executes an optimal trade-off between inverse filtering and noise smoothing. It removes the additive noise and inverts the blurring simultaneously.

Pseudo-inverse filter is a modified version of the inverse filter. It is actually a stabilized inverse filter. It gives much better results in comparison to inverse filter. It is too sensitive to noise.

9.8.3 Wiener Filter

Wiener filter is the most popular filter for restoration. It is not sensitive to noise. It exploits statistical properties of the image and can be used to restore the images in the presence of blur as well as noise.

The Wiener filtering is optimal in terms of the mean square error. In other words, it minimizes the overall mean square error in the process of inverse filtering and noise smoothing . The Wiener filtering is a linear estimation of the original image. The orthogonality principle implies that the Wiener filter in Fourier domain can be expressed as follows:

$$w(f1,f2) = \frac{H + (f1,f2) Sxx(f1,f2)}{|Hf1,f2|^2 Sxx(f1,f2) + S\eta\eta(f1,f2)}$$

Where, S_{xx} (f1, f2), (f1,f2) are respectively power spectra of the original image and the additive noise, and is the blurring filter.

Wiener filter has two separate parts, an inverse filtering part and a noise smoothing part. It not only performs the deconvolution by inverse filtering (high pass filtering) but also removes the noise with a compression operation (low pass filtering).

The advantages of Wiener filter are:

- Not sensitive to noise.
- Exploits the statistical properties of the image.

- Restores images in the presence of blur as well as noise

The disadvantage is that prior knowledge of the power spectral density of the original image is required, which is unavailable in practice.

9.8.4 Non Linear Image Restoration Techniques

In recent years, a variety of non-linear median type filters such as rank conditioned, eighted median, relaxed median, rank selection have been developed to overcome the shortcoming of linear filter.

Non-Linear technique does not explicitly implement the inverse. It uses an iterative approach to produce a successive improvement to the restoration until a termination condition is reached. It can cope up with some missing frequency components, with non -Gaussian noise and non- negativity etc.

9.8.5 Iterative Method

Iterative algorithms are used for restoring images which are suffering from a variety of degradations such as, linear, nonlinear, spatially 66 invariant blurs and signal-dependent noise, because of the flexible framework provided by each approach.

The basic form of iterative restoration filters is the one that iteratively approaches the solution of the inverse filter.

Advantages of the iterative method are:

- The first advantage of iterative method is that it does not require the convolution of images with 2D point spread function (PSF) containing many coefficients. The only convolution is that of the restored image with the PSF of the blur, which has relatively few coefficients.
- The second advantage is that, although the iteration produces the inverse filtered image as a result, if the iteration is continued indefinitely, the iteration can be terminated

whenever an acceptable restoration result has been successfully achieved.

- The third advantage is: no Fourier transforms are required, as it applied to images of arbitrary size.
- The fourth advantage is: these iteration methods can be easily extended for spatially variant restoration, i.e. the restoration where either the PSF of the blur or models of the image vary locally..

The drawbacks of iterative image restoration techniques is that they require more computational time than other techniques.

9.8.6 Maximum Likelihood Method

In case the point-spread function does not have characteristic spectral zeros or in case a parametric blur model such as motion or out-of-focus blur cannot be assumed, the individual coefficients of the point-spread function have to be estimated. To this end, maximum likelihood estimation procedures for the unknown coefficients have been developed.

Maximum likelihood estimation is a well-known technique for parameter estimation in situations where no stochastic knowledge is available about the parameters to be estimated. Most maximum likelihood identification techniques begin by assuming that the ideal image can described with the 2-D auto-regressive model. The parameters of this image model are not necessarily assumed to be known.

9.9 STOCHASTIC IMAGE RESTORATION TECHNIQUE

A Stochastic approach can be used for computing the maximum a posteriori (MAP) estimate of the original image given the degraded image. A stochastic relaxation algorithm generates a sequence of images that converges in an appropriate sense to the MAP estimate.

This sequence evolves by local (and potentially parallel) changes

in pixel gray levels and in locations and orientations of boundary elements. Deterministic, iterative-improvement methods generate a sequence of images that monotonically increase the posterior distribution (our “objective function”). In contrast, stochastic relaxation permits changes that decrease the posterior distribution as well. These are made on a random basis, the effect of which is to avoid convergence to local maxima.

The stochastic relaxation algorithm can be informally described as follows.

- A local change is made in the image based upon the current values of pixels and boundary elements in the immediate “neighborhood.” This change is random, and is generated by sampling from a local conditional probability distribution.
- The local conditional distributions are dependent on a global control parameter T called “temperature.” At low temperatures the local conditional distributions concentrate on states that increase the objective function, whereas at high temperatures the distribution is essentially uniform. The limiting cases, T= 0 and T= “, correspond respectively to greedy algorithms (such as gradient ascent) and undirected (i.e., “purely random”) changes. (High temperatures induce a loose coupling between neighboring pixels and a chaotic appearance to the image. At low temperatures the coupling is tighter and the images appear more regular.)
- The image restorations avoid local maxima by beginning at high temperatures where many of the stochastic changes will actually decrease the objective function. As the relaxation proceeds, temperature is gradually lowered and the process behaves increasingly like iterative improvement. (This gradual reduction of temperature simulates “annealing,” a procedure by which certain chemical systems can be driven to their low energy, highly regular, states.)

CHECK YOUR PROGRESS

Q 6: Which values does Gaussian noise generally disturb in digital images.

.....

.....

Q 7: Brownian noise can also be generated from white noise. (True/False)?

.....

.....

Q 8: Which method of image restoration can be employed if the prior knowledge about degradation is known?

.....

.....

Q 9: What are the two separate parts of Wiener filter?

.....

.....



Q 10: Most maximum likelihood identification techniques begin by assuming that the ideal image can described with the …… model. (Fill in the blank)



9.10 LET US SUM UP

- Image Restoration is the operation of taking a corrupt/noisy image and estimating the clean, original image. Corruption may come in many forms such as motion blur, noise and camera mis-focus.
- Image restoration is performed by reversing the process that blurred the image and this is performed by imaging a point source and using the point source image, which is called the Point Spread Function (PSF) to restore the image information lost to the blurring process.
- Blur is when the edge content of an image is reduced and the transition from one color to the other is very smooth. There are

different types of blurs - Defocus aberration, Motion blur, Bokeh, Box blur, and Gaussian blur.

- Noise is normally a high frequency component and a low pass filtering technique is used to remove the noise.
- Some Noise Models are Gaussian Noise model, White noise, Rayleigh noise, etc.
- If the prior knowledge about degradation is known then the deterministic method of image restoration can be employed. Deterministic techniques are of two types – linear and non-linear.
- The most common, simplest and the fastest kind of filtering is achieved by linear filters. The linear filter replaces each pixel with a linear combination of its neighbors and convolution kernel is used in prescription for the linear combination.
- Inverse Filter helps to get rid of the blurring (Deconvolution). The inverse filtering is a restoration technique for deconvolution, i.e., when the image is blurred by a known low pass filter, it is possible to recover the image by inverse filtering or generalized inverse filtering.
- Pseudo-inverse filter is a modified version of the inverse filter. It gives much better results in comparison to inverse filter.
- Wiener filter is the most popular filter for restoration. It is not sensitive to noise. It exploits statistical properties of the image and can be used to restore the images in the presence of blur as well as noise.
- Non -Linear techniques use an iterative approach to produce a successive improvement to the restoration until a termination condition is reached.
- Iterative algorithms are used for restoring images which are suffering from a variety of degradations.
- In case the point-spread function does not have characteristic spectral zeros or in case a parametric blur model such as motion or out-of-focus blur cannot be assumed, the individual coefficients of the point-spread function have to be estimated. To this end, maximum likelihood estimation procedures for the unknown coefficients have been developed.

- A Stochastic approach can be used for computing the maximum a posteriori (MAP) estimate of the original image given the degraded image.



9.11 FURTHER READIN

1) Gunturk, B. K., & Li, X. (Eds.). (2012). Image restoration: fundamentals and advances. CRC Press.
2) Katsaggelos, A. K. (2012), Digital Image Restoration (Springer Series in Information Sciences), Springer
3) Salazar, N. (2015). Developments and Applications of Image Restoration. Clanrye International.

9.12 ANSWERS TO CHECK YOUR PROGRESS



Ans to Q No 1: False

Ans to Q No 2: To reduce noise and recover resolution loss.

Ans to Q No 3: Additive

Ans to Q No 4: Defocus aberration

Ans to Q No 5: To reduce image noise and reduce detail

Ans to Q No 6: Gray values

Ans to Q No 7: True

Ans to Q No 8: Deterministic method

Ans to Q No 9: An inverse filtering part and a noise smoothing part

Ans to Q No 10: 2-D auto-regressive



9.13 MODEL QUESTIONS

Q 1: What is image degradation and restoration?

Q 2: What is image blur? What are the different types of common image blur?

Q 3: What is noise? What are the main sources of noise?

Q 4: What are the different types of noise models?

Q 5: What are the various types of Image Restoration Techniques?

*** ***** ***

UNIT 10 : IMAGE NOISE

UNIT STRUCTURE

10.1 Learning Objectives

10.2 Introduction

10.3 Classification of Image Noise

10.4 Adaptive Noise

10.5 Multiplicative Noise

10.6 Impulse Noise

10.7 Median Filtering

10.8 Trimmed Average Filter

10.9 Let Us Sum Up

10.10 Further Readings

10.11 Answers to Check Your Progress

10.12 Model Questions

10.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- learn about image noise
- describe different types of noise
- define adaptive, multiplicative and impulse noise
- learn about noise filtering

10.2 INTRODUCTION

Image noise is a random variation of brightness or color information in images, and is usually an aspect of electronic noise. It can be produced by the sensor and circuitry of a scanner or digital camera. Image noise can also originate in film grain and in the unavoidable shot noise of an ideal photon detector. Image noise is an undesirable by-product of image capture that obscures the desired information.

The original meaning of “noise” was “unwanted signal”; unwanted electrical fluctuations in signals received by AM radios causing audible

acoustic noise (“static”). By analogy, unwanted electrical fluctuations are also called “noise”.

Image noise can range from almost imperceptible specks on a digital photograph taken in good light to optical and radio-astronomical images that are almost entirely noise from which a small amount of information can be derived by sophisticated processing. Such a noise level would be unacceptable in a photograph since it would be impossible even to determine the subject.

Noise tells unwanted information in digital images. Noise produces undesirable effects such as artifacts, unrealistic edges, unseen lines, corners, blurred objects and disturbs the background scenes. To reduce these undesirable effects, prior learning of noise models is essential for further processing

10.3 CLASSIFICATION OF IMAGE NOISE

Digital noise may arise from various kinds of sources such as Charge Coupled Device (CCD) and Complementary Metal Oxide Semiconductor (CMOS) Sensors. In some sense, Points Spreading Function (PSF) and Modulation Transfer Function (MTF) have been used for timely, complete and quantitative analysis of noise models. Probability Density Function (PDF) or Histogram is also used to design and characterize the noise models.

Noise can generally be grouped into two classes:

- Image-independent noise
- Image-dependent noise

There are different types of noise that may be present in an image. Some common types are listed below:

- Image independent noise can often be described by an additive noise model, where the recorded image f(i,j) is the sum of the true image s(i,j) and the noise n(i,j):

$$f(i,j) = s(i,j) + n(i,j)$$

The noise n(i,j) is often zero-mean and described by its

variance σ_n^2. The impact of the noise on the image is often described by the signal to noise ratio (SNR), which is given by

$$SNR = \frac{\sigma_s}{\sigma_n} = \frac{\sigma_f^2}{\sigma_n^2}$$

where σ_n^2 and are the variances of the true image and the recorded image, respectively.

- In the second case of *data-dependent noise* (e.g. arising when monochromatic radiation is scattered from a surface whose roughness is of the order of a wavelength, causing wave interference which results in image speckle), it is possible to model noise with a multiplicative, or non-linear, model. These models are mathematically more complicated; hence, if possible, the noise is assumed to be data independent.



CHECK YOUR PROGRESS

Q 1: Image noise is usually an aspect of electronic noise. True/False?

Q 2: What are the two classes of image noise?

.....

.....

Q 3: Which image model is mathematically more complicated?

.....

.....

Q 4: What is SNR?

.....

.....

10.4 ADAPTIVE NOISE

It is also called Gaussian noise or electronic noise because it arises in amplifiers or detectors. It is caused by natural sources such as thermal vibration of atoms and discrete nature of radiation of warm

objects. This noise generally disturbs the gray values in digital images. That is why this noise model is essentially designed and characterized by its Probability Distribution Function (PDF) or normalizes histogram with respect to gray value.

This is given as:

$$P(g) = \sqrt{\frac{1}{2\pi\sigma^2}} e^{-\frac{(g-\mu)^2}{2\sigma}}$$

where g = gray value, *s*= standard deviation and *m*= mean.

Generally Gaussian noise mathematical model represents the correct approximation of real world scenarios. In this noise model, the mean value is zero, variance is 0.1 and 256 gray levels in terms of its PDF.

10.5 MULTIPLICATIVE NOISE

In signal processing, the term multiplicative noise refers to an unwanted random signal that gets multiplied into some relevant signal during capture, transmission, or other processing.

An important example is the speckle noise commonly observed in radar imagery. Examples of multiplicative noise affecting digital photographs are proper shadows due to undulations on the surface of the imaged objects, shadows cast by complex objects like foliage and Venetian blinds, dark spots caused by dust in the lens or image sensor, and variations in the gain of individual elements of the image sensor array.

Speckle is a granular interference that inherently exists in and degrades the quality of the active radar, synthetic aperture radar (SAR), medical ultrasound and optical coherence tomography images.

The vast majority of surfaces, synthetic or natural, are extremely rough on the scale of the wavelength. Images obtained from these surfaces by coherent imaging systems such as laser, SAR, and ultrasound suffer from a common interference phenomenon called speckle. The origin of this phenomenon is seen if we model our reflectivity function as an array of scatterers. Because of the finite resolution, at any time we are receiving from a distribution of scatterers within the resolution cell.

These scattered signals add coherently; that is, they add constructively and destructively depending on the relative phases of each scattered waveform. Speckle results from these patterns of constructive and destructive interference shown as bright and dark dots in the image.

Although commonly referred to as “speckle noise”, speckle is not noise in its generally understood sense of an unwanted modification to a desired signal. Rather, it is the signal itself that fluctuates, because the scatterers are not identical for each cell, and the signal is highly sensitive to small variations in scatterers.

10.6 IMPULSE NOISE

Impulse noise corruption is very common in digital images. Impulse noise is always independent anduncorrelated to the image pixels and is randomly distributed over the image. For an impulse noise corrupted image all the image pixels are not noisy, a number of image pixels will be noisy andthe rest of pixels will be noise free.

There are two types of impulse noise:

- **Salt and pepper type of noise:** In salt and pepper type of noise the noisy pixels takes either salt value (gray level -225) or pepper value (grey level -0) and it appears as black and white spots on the images. If p is the total noise density then salt noise and pepper noise will have a noise density of p/2.
- **Random valued impulse noise:** In case of random valued impulse noise, noise can take any gray level value from zero to 225. In this case also the noise is randomly distributed over the entire image and probability of occurrence of any gray level value as the noise will be same.

10.7 MEDIAN FILTERING

The Median Filter is a non-linear digital filtering technique, often used to remove noise from an image or signal. Such noise reduction is a typical pre-processing step to improve the results of later processing

(for example, edge detection on an image). Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise (but see discussion below). It also has applications in signal processing.

The main idea of the median filter is to run through the signal entry by entry, replacing each entry with the median of neighboring entries. The pattern of neighbors is called the “window”, which slides, entry by entry, over the entire signal. For 1D signals, the most obvious window is just the first few preceding and following entries, whereas for 2D (or higher-dimensional) signals such as images, more complex window patterns are possible (such as “box” or “cross” patterns). Note that if the window has an odd number of entries, then the median is simple to define: it is just the middle value after all the entries in the window are sorted numerically. For an even number of entries, there is more than one possible median.

The median filter considers each pixel in the image in turn and looks at its nearby neighbors to decide whether or not it is representative of its surroundings. Instead of simply replacing the pixel value with the mean of neighboring pixel values, it replaces it with the median of those values. The median is calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value. (If the neighborhood under consideration contains an even number of pixels, the average of the two middle pixel values is used.)

10.8 TRIMMED AVERAGE FILTER

It is a non-linear smoothing filter.and is also called the á-trimmed mean filter.

Truncating or ‘trimming’ the distribution before taking the mean, by removing some proportion (usually called á) of the largest and smallest values, is a simple way of ensuring that extreme local values do not influence the output. The portion to be truncated varies between 0% (equivalent to the mean) and 100% (equivalent to the median).

In order to calculate the á-trimmed mean, the data should be sorted low to high and summed the central part of the ordered array. The number of data values which are dropped from the average is controlled by the trimming parameter á which assumes values between 0 and 0.5.

In the case that á is zero, none of the data samples will be rejected and the summation will be computed over the entire samples. On the other hand, if á is opted 0.5, all samples, except for the middle one, are rejected and the á-trimmed mean will turn into median filter.



CHECK YOUR PROGRESS

Q 5: How is Gaussian noise caused?

.....

.....

Q 6: Where is speckle noise commonly encountered?

.....

.....

Q 7: Impulse noise is always image-dependent. True/False?

.....

.....

Q 8: What are the two types of impulse noise?

.....

.....

Q 9: What kind of filtering technique is Median filter?

.....

.....



10.9 LET US SUM UP

- Image noise is a random variation of brightness or color information in images, and is usually an aspect of electronic noise. It can be produced by the sensor and circuitry of a scanner or digital camera.
- Digital noise may arise from various kinds of sources such as

Charge Coupled Device (CCD) and Complementary Metal Oxide Semiconductor (CMOS) sensors.

- Noise can generally be grouped into two classes: Image-independent noise, and Image-dependent noise
- Adaptive noise is also called as Gaussian noise or electronic noise because it arises in amplifiers or detectors. It is caused by natural sources such as thermal vibration of atoms and discrete nature of radiation of warm objects. This noise generally disturbs the gray values in digital images.
- Multiplicative noise refers to an unwanted random signal that gets multiplied into some relevant signal during capture, transmission, or other processing.
- Impulse noise is always independent and uncorrelated to the image pixels and is randomly distributed over the image. For an impulse noise corrupted image all the image pixels are not noisy, a number of image pixels will be noisy and the rest of pixels will be noise free.
- The median filter is a non-linear digital filtering technique, often used to remove noise from an image or signal.
- Truncating, or 'trimming', the distribution before taking the mean, by removing some proportion (usually called ?) of the largest and smallest values, is a simple way of ensuring that extreme local values do not influence the output.



10.10 FURTHER READING

1. Gunturk, B. K., & Li, X. (Eds.). (2012). Image restoration: fundamentals and advances. CRC Press.
2. Gonzales, R. C., & Woods, R. E. (2008). Digital image processing. Prentice Hall.
3. Jayaraman, S. Esakkirajandan, S., & Veerakumar, T. (2009). Digital Image Processing, New Delhi: Tata McGraw-Hill Education Private Limited.

10.11 ANSWERS TO CHECK YOUR PROGRESS

Ans to Q No 1: True

Ans to Q No 2: Image-independent and image-dependent

Ans to Q No 3: Image-dependent

Ans to Q No 4: Signal to noise Ratio

Ans to Q No 5: By natural sources

Ans to Q No 6: Radar imagery

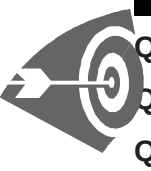
Ans to Q No 7: False

Ans to Q No 8: Salt and pepper, and random-valued

Ans to Q No 9: Non-linear digital filtering technique



10.12 MODEL QUESTIONS



Q 1: What is image noise? Describe the two classes of image noise.

Q 2: What is adaptive noise?

Q 3: How does multiplicative noise affect an image?

Q 4: What is impulse noise? What are the different types of impulse noise?

Q 5: How is median filtering used to reduce image noise?

Q 6: What are the characteristics of Trimmed Average Filter?

*** ***** ***

UNIT 11: IMAGE SEGMENTATION

UNIT STRUCTURE

11.1 Learning Objectives

11.2 Introduction

11.3 Classification of Image Segmentation Techniques

11.4 Local and Global Segmentation

11.5 Image Segmentation Based on Thresholding

11.6 Region Approach to Image Segmentation

- 11.6.1 Region Growing Methods
- 11.6.2 Region Splitting and Merging Methods

11.7 Clustering Technique

11.8 Edge Detection Techniques

11.9 Edge Linking

- 11.9.1 Hough Transform

11.10 Watershed Transformation

11.11 Let Us Sum Up

11.12 Further Reading

11.13 Answers to Check Your Progress

11.14 Model Questions

11.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- learn about image segregation
- learn about the various Image segmentation techniques
- learn about Local and global segmentation

11.2 INTRODUCTION

Image segmentation is an important and challenging process of image processing. Image segmentation technique is used to partition an image into meaningful parts having similar features and properties. The

main aim of segmentation is simplification i.e. representing an image into meaningful and easily analyzable way. Image segmentation is the first step in image analysis. The goal of image segmentation is to divide an image into several parts/segments having similar features or attributes. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image so that pixels with the same label share certain characteristics.The basic applications of image segmentation are: Content-based image retrieval, Medical imaging, Object detection and Recognition Tasks, Automatic traffic control systems and Video surveillance, etc.

11.3 CLASSIFICATION OF IMAGE SEGMENTATION TECHNIQUES

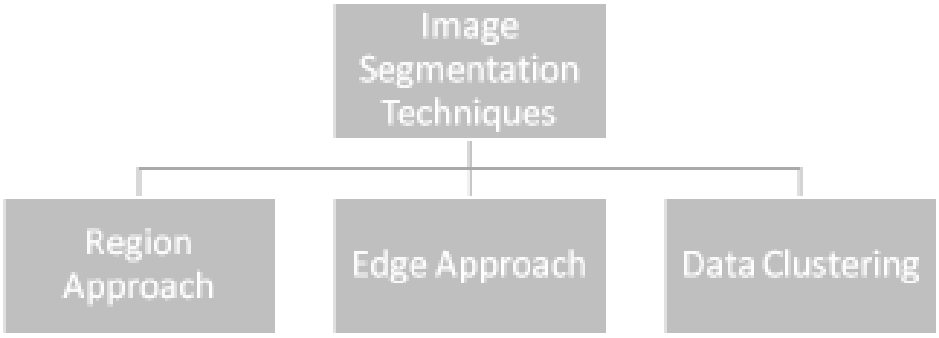
The image segmentation approaches can be categorized into two types based on properties of image.

- **Discontinuity detection based approach:** This is the approach in which an image is segmented into regions based on discontinuity. The edge detection based segmentation falls in this category in which edges formed due to intensity discontinuity are detected and linked to form boundaries of regions.
- **Similarity detection based approach:** This is the approach in which an image is segmented into regions based on similarity. The techniques that falls under this approach are: thresholding techniques, region growing techniques and region splitting and merging. These all divide the image into regions having similar set of pixels. The clustering techniques also use this methodology. They divide the image into set of clusters having similar features based on some predefined criteria

In other words, we can say that image segmentation can be approached from three perspectives: Region approach, Edge approach and Data clustering. The region approach falls under similarity detection while edge detection and boundary detection falls under discontinuity

detection. Clustering techniques are also under similarity detection.

Figure 11.1: Image Segmentation Approaches



11.4 LOCAL AND GLOBAL SEGMENTATION

The image segmentation can be classified into two basic types: local segmentation (concerned with specific part or region of image) and global segmentation (concerned with segmenting the whole image, consisting of large number of pixels).

The global properties of the image are characterized by the mean values of different pixel classes and the continuous boundary of the object or region. The local properties are characterized by the interactions of neighboring pixels and the image edge.

11.5 IMAGE SEGMENTATION BASED ON THRES HOLDING

Thresholding methods are the simplest methods for image segmentation. These methods divide the image pixels with respect to their intensity level. These methods are used over images having lighter objects than the background. The selection of these methods can be manual or automatic i.e. can be based on prior knowledge or information of image features.

There are basically three types of thresholding:

- **Global Thresholding:** This is done by using any appropriate threshold value/T. This value of T will be constant for the whole image. On the basis of T, the output image *q(x,y)* can be obtained from the original image *p(x,y)* as:

q(x,y) = 1, if p(x,y) > T

q(x,y) = 0, if p(x,y) <= T

- **Variable Thresholding:** In this type of thresholding, the value of T can vary over the image. This can further be of two types:
  - Local Threshold: In this the value of T depends upon the neighborhood of x and y.
  - Adaptive Threshold: The value of T is a function of x and y.
- **Multiple Thresholding:** In this type of thresholding, there are multiple threshold values like T0 and T1. By using these, the output image can be computed as:

q(x,y) = m, if p(x,y) > T1

q(x,y) = n, if p(x,y) <= T1

q(x,y) = o, if p(x,y) <= T0

The values of thresholds can be computed with the help of the peaks of the image histograms. Simple algorithms can also be generated to compute these.



CHECK YOUR PROGRESS

Q 1: The main aim of segmentation is _____.

.....

.....

Q 2: What are the two categories of image segmentation approaches?

.....

.....

Q 3: Thresholding methods divide the image pixels with respect to their_____ level.

Q 4: What are the three types of thresholding?

.....

.....

Q 5: How can the values of thresholds can be computed?

.....

.....

11.6 REGION APPROACH TO IMAGE SEGMENTATION

The region based segmentation methods are the methods that segments the image into various regions having similar characteristics. There are two basic techniques based on this method:

- Region growing method
- Region splitting and merging method

11.6.1 Region Growing Methods

The region growing based segmentation methods are the methods that segments the image into various regions based on the growing of seeds (initial pixels). These seeds can be selected manually (based on prior knowledge) or automatically (based on particular application). Then the growing of seeds is controlled by connectivity between pixels and with the help of the prior knowledge of problem, this can be stopped.

The basic algorithm (based on 8- connectivity) steps for region growing method are:

If *p(x,y)* is the original image that is to be segmented and *s(x,y)* is the binary image where the seeds are located. Let T be any predicate which is to be tested for each *(x,y)* location.

- First of all, all the connected components of s are eroded.
- Compute a binary image PT. Where P_T (x, y) = 1, if T(x, y) = True.
- Compute a binary image q, where q(x, y) = 1, if P_T(x, y) = 1 and (x, y) is 8-connected to seed in s.

These connected components in q are segmented regions.

11.6.2 RegionSplitting and Merging Methods

The region splitting and merging based segmentation methods use two basic techniques i.e. splitting and merging for segmenting an image into various regions. Splitting stands for iteratively dividing an image into regions having similar

characteristics and merging contributes to combining the adjacent similar regions.

The basic algorithm steps for region growing and merging are:

Let p be the original image and T be the particular predicate.

- First of all the R_1 is equal to p.
- Each region is divided into quadrants for which T (R_i) = False.
- If for every region, T (R_j) = True, then merge adjacent regions R_i and R_j such that T (R_i U R_j) = True.
- Repeat step 3 until merging is impossible.

11.7 CLUSTERING TECHNIQUE

The clustering based techniques are the techniques, which segment the image into clusters having pixels with similar characteristics. Data clustering is the method that divides the data elements into clusters such that elements in the same cluster are similar to each other than others. There are two basic categories of clustering methods: Hierarchical method and Partition based method. The hierarchical methods are based on the concept of trees. In this the root of the tree represents the whole database and the internal nodes represent the clusters. On the other side the partition based methods use optimization methods iteratively to minimize an objective function. In between these two methods there are various algorithms to find clusters.

There are basic two types of clustering:

- **Hard Clustering:** Hard clustering is a simple clustering technique that divides the image into set of clusters so that one pixel can only belong to only one cluster. In other words, it can be said that each pixel can belong to exactly one cluster. These methods use membership functions having values either 1 or 0 i.e., one certain pixel can belong to either a particular cluster or not. An example of a hard clustering based technique is one k-means clustering based technique known as HCM. In this technique, first of all the centers are computed then each pixel is assigned to nearest center. It emphasizes on maximizing the intra cluster similarity and also

minimizing the inter cluster equality.

- **Soft Clustering:** The soft clustering is a more natural type of clustering because in real life exact division is not possible due to the presence of noise. Thus, soft clustering techniques are most useful for image segmentation in which division is not strict. The example of such type of technique is fuzzy c-means clustering. In this technique pixels are partitioned into clusters based on partial membership i.e. one pixel can belong to more than one cluster and this degree of belonging is described by membership values. This technique is more flexible than other techniques.

11.8 EDGE DETECTION TECHNIQUES

The edge detection techniques are well developed techniques of image processing on their own. The edge based segmentation methods are based on the rapid change of intensity value in an image because a single intensity value does not provide good information about edges. Edge detection techniques locate the edges-where either the first derivative of intensity is greater than a particular threshold or the second derivative has zero crossings. In edge-based segmentation methods, first of all the edges are detected and then they are connected together to form the object boundaries to segment the required regions. The basic two edge based segmentation methods are: Gray histograms and Gradient based methods. To detect the edges one of the basic edge detection techniques like sobel operator, canny operator and Robert s operator etc. can be used. Result of these methods is basically a binary image.

Pfaltz (1966) gave an efficient algorithm for doing this for 4- and 8-connected regions, termed a connected components algorithm.

The algorithm operates on a raster scan, in which each pixel is visited in turn, starting at the top-left corner of the image and scanning along each row, finishing at the bottom-right corner. For each non-edge pixel, (i, j), the following conditions are checked.

If its already visited neighbours — (i “ 1, j) and (i, j “ 1) in the 4-

connected case, also (i “ 1, j “ 1) and (i “ 1, j + 1) in the 8-connected case — are all edge pixels, then a new category is created and (i, j) is allocated to it. Alternatively, if all its non-edge neighbours are in a single category, then (i, j) is also placed in that category. The final possibility is that neighbours belong to two or more categories, in which case (i, j) is allocated to one of them and a note is kept that these categories are connected and therefore, should from then on be considered as a single category.

11.9 EDGE LINKING

Edge detectors yield pixels in an image lie on edges.

The next step is to try to collect these pixels together into a set of edges.

Thus, our aim is to replace many points on edges with a few edges themselves.

The practical problem may be much more difficult than the idealized case.

- Small pieces of edges may be missing,
- Small edge segments may appear to be present due to noise where there is no real edge, etc.

In general, edge linking methods can be classified into two categories:

- **Local Edge Linkers**, where edge points are grouped to form edges by considering each point’s relationship to any neighbouring edge points. Local edge linking methods usually start at some arbitrary edge point and consider points in a local neighbourhood for similarity of edge direction.
  - If the points satisfy the similarity constraint then the points are added to the current edge set. The neighbourhoods based around the recently added edge points are then considered in turn and so on.
  - If the points do not satisfy the constraint then we conclude that we are at the end of the edge, and so the process stops. A new starting edge point is found which does not belong to any edge set found so far, and the process is repeated.

- The algorithm terminates when all edge points have been linked to one edge or at least have been considered for linking once.

Thus, the basic process used by local edge linkers is that of tracking a sequence of edge points. An advantage of such methods is that they can readily be used to find arbitrary curves.

- **Global Edge Linkers**, where all edge points in the image plane are considered at the same time and sets of edge points are sought according to some similarity constraint, such as points which share the same edge equation.

11.9.1 Hough Transform

One powerful global method for detecting edges is called the Hough Transform.

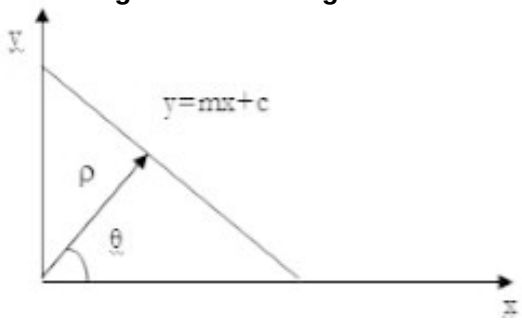
Let us suppose that we are looking for straight lines in an image.

The straight line represented by y=mx+c can be expressed in polar coordinate system as,

$\rho = x\cos(\theta) + y\sin(\theta)$ ……..(i)

Where defines a vector from the origin to the nearest pointon the straight line y=mx+c. this vector will be perpendicular from the origin to the nearest point to the line as shown in the below figure.

Figure 11.2: Straight LIne



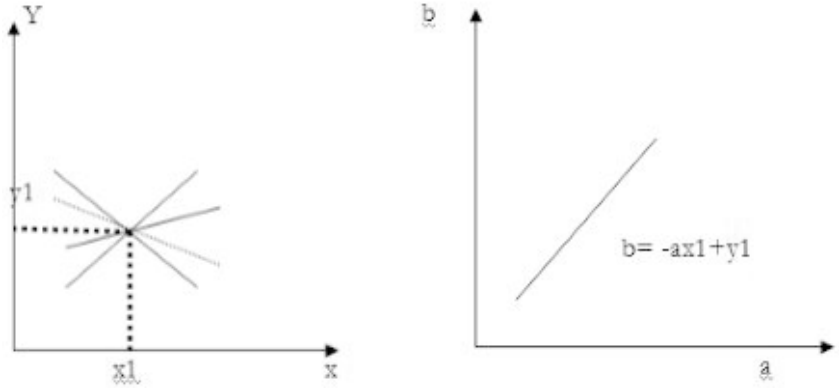
Any line in the x, y plane corresponds to the point in the 2D space defined by the parameter and ,θ . This the Hough transform of a straight line in the x,y plane is a single point in the space and these points should satisfy the given equation with x1,y1 as constants. Thus the locus of all such lines in the x, y plane

corresponds to the particular sinusoidal curve in the ρ,θ space.

Suppose we have the edge points xi,yi that lie along the straight line having parameters $\rho 0,\theta 0$. Each edge point plots to a sinusoidal curve in the ñ,è space, but these curves must intersect at a point $\rho 0,\theta 0$. Since this is a line they all have in common. For example considering the equation of a line: y1= ax1+b

Using this equation and varying the values of a and b, infinite lines can pass through this point (x1,y1).

Figure 11.3



However, if we write the equation as

B= -ax1+y1

and then consider the ab plane instead of xy plane, we get a straight line for a point (xi,yi). This entire line in the ab plane is due to a single point in the xy plane and different values of and b. Now consider another point (x2, y2) in the xy plane. The slope intercept equation of this line is,

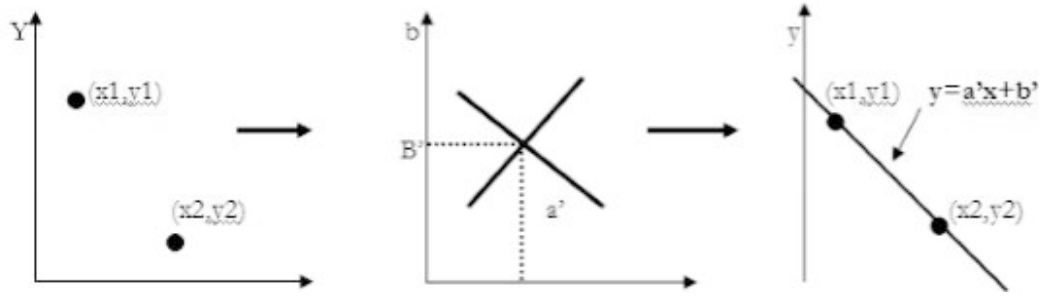
Y2= ax2+ b……(1)

Writing the equation in terms of the ab plane we get,

B= -ax2+y2……..(2)

This is another line in the ab plane. These two line will intersect each other somewhere in the ab plane only if they are part of a straight line in the xy plane. The point of intersection in the ab plane is noted as (a',b'). using this (a',b') in the standard slope-intercept form i.e. y=a'x+b', we get a line that passes through the points (x1 , y1) and (x2, y2) in the xy plane.

Figure 11.4



One useful property of the Hough transform is that the pixels which lie on the line need not all be contiguous.

11.10 WATERSHED TRANSFORMATION

A watershed is a transformation defined on a grayscale image. The name refers metaphorically to a geological watershed, or drainage divide, which separates adjacent drainage basins. The watershed transformation treats the image it operates upon like a topographic map, with the brightness of each point representing its height, and finds the lines that run along the tops of ridges.

There are different technical definitions of a watershed. In graphs, watershed lines may be defined on the nodes, on the edges, or hybrid lines on both nodes and edges. Watersheds may also be defined in the continuous domain.

There are also many different algorithms to compute watersheds.

- **Watershed by flooding**

  The idea was introduced in 1979 by S. Beucher and C. Lantuéjoul.[2] The basic idea consisted of placing a water source in each regional minimum in the relief, to flood the entire relief from sources, and build barriers when different water sources meet. The resulting set of barriers constitutes a watershed by flooding. A number of improvements, collectively called Priority-Flood, have since been made to this algorithm,

- **Watershed by topographic distance**

  Intuitively, a drop of water falling on a topographic relief flows towards the “nearest” minimum. The “nearest” minimum is that minimum which

lies at the end of the path of steepest descent. In terms of topography, this occurs if the point lies in the catchment basin of that minimum. The previous definition does not verify this condition.

- **Watershed by the drop of water principle**

  Intuitively, the watershed is a separation of the regional minima from which a drop of water can flow down towards distinct minima. A formalization of this intuitive idea was provided in for defining a watershed of an edge-weighted graph.

- **Inter-pixel watershed**

  S. Beucher and F. Meyer introduced an algorithmic inter-pixel implementation of the watershed method, given the following procedure:

  - Label each minimum with a distinct label. Initialize a set S with the labeled nodes.
  - Extract from S a node x of minimal altitude F, that is to say F(x) = min{F(y)|y " S}. Attribute the label of x to each non-labeled node y adjacent to x, and insert y in S.
  - Repeat Step 2 until S is empty.

  The non-labeled pixels are the watershed lines.



CHECK YOUR PROGRESS

Q 6: What are seeds?

.....

.....

Q 7: What are the two types of region based techniques?

.....

.....

Q 8: What are the two types of clustering?

.....

.....

Q 9: What is the result of an edge detection technique?

.....

.....

Q 10: A watershed is a transformation defined on a _____ image.

.....

.....



11.11 LET US SUM UP

- Image segmentation technique is used to partition an image into meaningful parts having similar features and properties.
- The goal of image segmentation is to divide an image into several parts/segments having similar features or attributes.
- The image segmentation approaches can be categorized into two types based on properties of image - Discontinuity detection based approach & Similarity detection based approach.
- Thresholding methods divide the image pixels with respect to their intensity level.
- The region based segmentation methods are the methods that segments the image into various regions having the similar characteristics. There are two basic techniques based on this method: Region growing method & Region splitting and merging method
- The clustering based techniques are the techniques, which segment the image into clusters having pixels with similar characteristics. Data clustering is the method that divides the data elements into clusters such that elements in same cluster are more similar to each other than others.
- The edge based segmentation methods are based on the rapid change of intensity value in an image because a single intensity value does not provide good information about edges.
- A watershed is a transformation defined on a grayscale image.



11.12 FURTHER READING

1. Gonzales, R. C., & Woods, R. E. (2008). Digital image processing.

Prentice Hall.

2. Jayaraman, S. (2009). S. Esakkirajandan T. Veerakumar. Digital Image Processing, New Delhi: Tata McGraw-Hill Education Private Limited.

11.13 ANSWERS TO CHECK YOUR PROGRESS

Ans to Q No 1: simplification

Ans to Q No 2: Discontinuity detection and Similarity detection

Ans to Q No 3: intensity

Ans to Q No 4: Global, variable and multiple

Ans to Q No 5: Image histograms

Ans to Q No 6: Initial pixels

Ans to Q No 7: Region growing and Region splitting and merging.

Ans to Q No 8: Hard and soft

Ans to Q No 9: Binary Image

Ans to Q No 10: grayscale



11.14 MODEL QUESTIONS

Q 1: What is image segmentation? What are the different types of image segmentation techniques?

Q 2: What are local and global segmentations?

Q 3: Describe the different types of Thresholding.

Q 4: Describe the different region based techniques.

Q 5: What is clustering? What are the different types of clustering?

Q 6: Define the Hough Transform with example.

*** ***** ***

UNIT 12 : IMAGE COMPRESSION

UNIT STRUCTURE

12.1 Learning Objectives

12.2 Introduction

12.3 Need for Image Compression

12.4 Redundancy in Images

12.5 Image Compression Scheme

12.6 Huffman Coding

12.7 Transform Base Coding

12.8 Vector Quantization

12.9 Let Us Sum Up

12.10 Further Reading

12.11 Answers to Check Your Progress

12.12 Model Questions

12.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- learn about image compression and its need
- explain lossless and lossy image compression
- describe Hoffman coding image compression technique

12.2 INTRODUCTION

In today's life, an enormous amount of data is stored, processed and transmitted digitally. This online information is graphical and pictorial in nature, the storage and communications requirements are immense. Compressing data is needed for practical and commercial interest prior to storage and transmission. Image compression is an application of data compression which encodes the original image with a few bits. The image compression is developed with the objective to reduce the size of the image and to store or transmit data in an efficient form. The main goal of storage system is to reduce the quantity of storage as much as

possible, and the decoded image displayed in the monitor can be same as the original image. There are two different compression methods like *lossy* and *lossless* compression. Under lossless compression we can use Huffman code to compress the image, and under lossy category, we can use transform base code and quantization methods. Compression can be achieved by reducing the redundancy. Redundancy may occur in coding, in interpixel or in psychovisual.

12.3 NEED FOR IMAGE COMPRESSION

High quality image data requires large amount of storage stape and transmission bandwidth. The current technology is sometime unable to handle it technically and economically. For example, an image of size 1024 pixel x 1024 pixel x 24 bit, without compression, would require 3 MB of storage and 7 minutes for transmission, utilizing a high speed, 64 Kbit/s, ISDN line. If the image is compressed at a 10:1 compression ratio, the storage requirement is reduced to 300 KB and the transmission time drops to under 6 seconds. Seven 1 MB images can be compressed and transferred to a floppy disk in less time than it takes to send one of the original files, uncompressed, over an AppleTalk network.

In a distributed environment large image files remain a major bottleneck within systems. Compression is an important component of the solutions available for creating file sizes of manageable and transmittable dimensions. Increasing the bandwidth is another method, but the cost sometimes makes this a less attractive solution. Platform portability and performance are important in the selection of the compression/decompression technique to be employed. The easiest way to reduce the size of the image file is to reduce the size of the image itself. By shrinking the size of the image, fewer pixels need to be stored and consequently the file will take less time to load.

12.4 REDUNDANCY IN IMAGES

Data redundancy is a primary issue in digital image compression. It is not an abstract concept but a mathematically quantifiable entity. If

n_1 and n_2 denote the number of information-carrying units in two datasets that represent the same information, the relative data redundancy R_D of the first dataset (the one characterized by n_1) is defined as

$$R_D = 1 - \frac{1}{C_R} \quad \text{.....(12.1)}$$

where C_R , commonly called the compression ratio, is

$$C_R = \frac{n_2}{n_2} \quad \text{.....(12.2)}$$

For the case $n_2 = n_1$, $C_R = 1$ and $R_D = 0$, indicating that (relative to the second dataset) the first representation of the information contains no redundant data. When $n_2 << n1$, C_R → " and R_D → 1, implying significant compression and highly redundant data. Finally, when $n_2 >> n_1$, C_R → 0 and R_D → ", indicating that the second dataset contains much more data than the original representation. This, of course, is the normally undesirable case of data expansion. In general, C_R and R_D lie in the open intervals (0,") and (-", 1), respectively. A practical compression ratio, such as 10 (or 10:1), means that the first dataset has 10 information carrying units (say, bits) for every 1 unit in the second or compressed dataset. The corresponding redundancy of 0.9 implies that 90% of the data in the first data set is redundant. In digital image compression, three basic data redundancies can be identified and exploited: coding redundancy, interpixel redundancy, and psychovisual redundancy. Data compression is achieved when one or more of these redundancies are reduced or eliminated. Let us explain them in detail.

- **Coding Redundancy:** Coding redundancy is associated with the representation of information. The information is represented in the form of codes. The Huffman code and arithmetic codes are some examples of codes

  Let us assume that a discrete random variable r_k in the interval [0, 1] represents the gray levels of an image and that each r_k occurs with probability p_r (r_k).

$$C_R = \frac{n_2}{n_2} \quad \text{.....(12.3)}$$

where k = 0,1,2,3,……, L-1

L is the number of gray levels, n_k is the number of times that the k^{th} gray level appears in the image, and n is the total number of pixels in the image. If the number of bits used to represent each value of r_k is 1 (r_k), then the average number of bits required to represent each pixel is

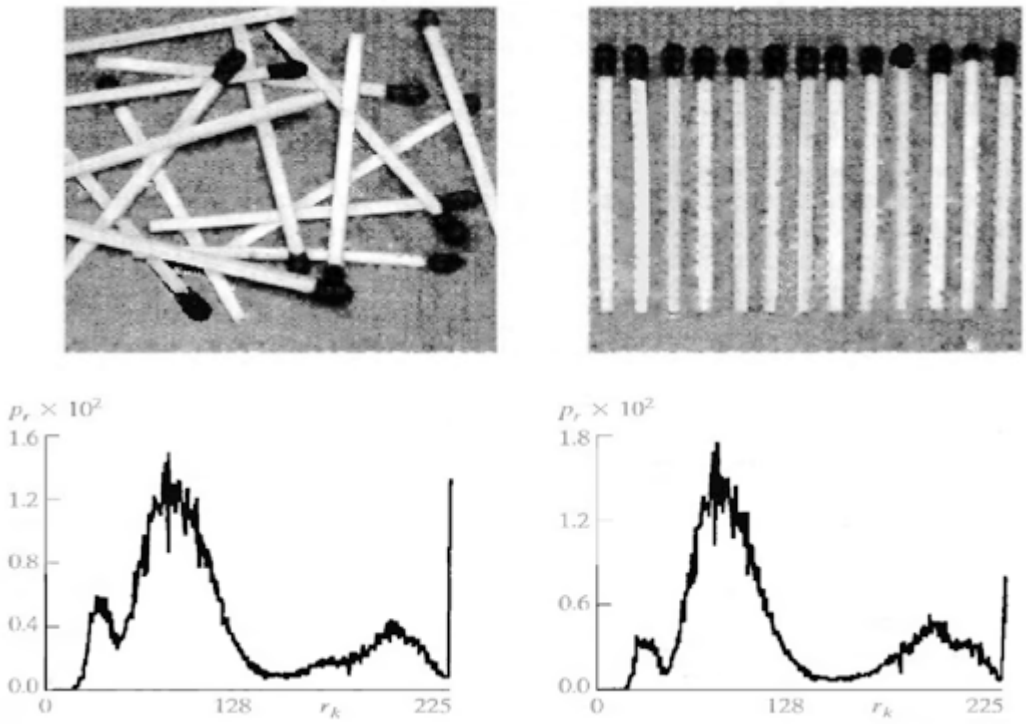
$$L_{avg} = \sum_{k=0}^{L-1} l(r_k)\, p(r_k) \quad \text{.....(12.4)}$$

That is, the average length of the code words assigned to the various gray-level values is found by summing up the product of the number of bits used to represent each gray level and the probability that the gray level occurs. Thus the total number of bits required to code an M X N image is MNL_{avg}.

- **Interpixel Redundancy:** Interpixel redundancy is due to the correlation between neighboring pixels in an image. It means that the neighbouring pixels are not statistically independent. The interpixel correlation is referred to as interpixel redundancy.

Let us consider the images shown in Fig. 12.1(a) and (b). Figure 12.1(c) and (d) shows that these images have virtually identical histograms. Both histograms are trimodal, indicating the presence of three dominant ranges of gray-level values. Because the gray levels in these images are not equally probable, variable-length coding can be used to reduce the coding redundancy that would result from a straight or natural binary encoding of their pixels. The coding process, however, would not alter the level of correlation between the pixels within the images. In other words, the codes used to represent the gray levels of each image have nothing to do with the correlation between pixels. These correlations result from the structural or geometric relationships between the objects in the image.

Fig.12.1 Two images and their gray-level histograms and normalized autocorrelation coefficients along one line.



$$\gamma(\Delta n) = \frac{A(\Delta n)}{A(0)} \quad \text{..... (12.5)}$$

where

$$A(\Delta n) = \frac{1}{N - \Delta n} \sum_{y=0}^{N-1-\Delta n} f(x,y)\, f(x, y+\Delta n) \quad \text{.....(12.6)}$$

The scaling factor in Eq. above accounts for the varying number of sum terms that arise for each integer value of Än. Of course, Än must be strictly less than N, the number of pixels on a line. The variable x is the coordinate of the line used in the computation. Note the dramatic difference between the shape of the functions shown in Figs. 12.1(e) and (f). Their shapes can be qualitatively related to the structure in the images in Figs. 12. 1 (a) and (b).This relationship is particularly noticeable in Fig. 12.1 (f), where the high correlation between pixels separated by 45 and 90 samples can be directly related to the spacing between the vertically oriented matches of Fig. 12.1(b). In addition, the adjacent pixels of both images are highly correlated.

When Än is 1, ã is 0.9922 and 0.9928 for the images of Figs. 12.1 (a) and (b), respectively. These values are typical of most properly sampled television images. These illustrations reflect another important form of data redundancy—one directly related to the interpixel correlations within an image. Because the value of any given pixel can be reasonably predicted from the value of its neighbors, the information carried by individual pixels is relatively small. Much of the visual contribution of a single pixel to an image is redundant; it could have been guessed on the basis of the values of its neighbors. A variety of names, including spatial redundancy, geometric redundancy, and interframe redundancy, have been coined to refer to these interpixel dependencies. We use the term interpixel redundancy to encompass them all. In order to reduce the interpixel redundancies in an image, the 2-D pixel array normally used for human viewing and interpretation m ust be transformed into a more efficient (but usually "non-visual") format. For example, the differences between adjacent pixels can be used to represent an image. Transformations of this type (that is, those that remove interpixel redundancy) are referred to as mappings. They are called reversible mappings if the original image elements can be reconstructed from the transformed data set.

- **Psychovisual Redundancy**: The brightness of a region, as perceived by the eye, depends on factors other than simply the light reflected by the region. For example, intensity variations (Mach bands) can be perceived in an area of constant intensity. Such phenomena result from the fact that the eye does not respond with equal sensitivity to all visual information. Certain information simply has less relative importance than other information in normal visual processing. This information is said to be psychovisually redundant. It can be eliminated without significantly impairing the quality of image perception.

  That psychovisual redundancies exist should not come as a surprise, because human perception of the information in an image

normally does not involve quantitative analysis of every pixel value in the image. In general, an observer searches for distinguishing features such as edges or textural regions and mentally combines them into recognizable groupings. The brain then correlates these groupings with prior knowledge in order to complete the image interpretation process. Unlike coding and interpixel redundancy, psychovisual redundancy is associated with real or quantifiable visual information. Its elimination is possible only because the information itself is not essential for normal visual processing. Since the elimination of psychovisually redundant data results in a loss of quantitative information, it is commonly referred to as quantization.

This terminology is consistent with normal usage of the word, which generally means the map mapping of a broad range of input values to a limited number of output values. As it is an irreversible operation (visual information is lost), quantization results in lossy data compression.

12.5 IMAGE COMPRESSION SCHEME

Image compression can be further classified or divided in two separate types such as lossy compression and lossless compression. In the lossy compression as its name indicated that it results in the loss of little information. In this technique the compressed image is same as to actual/original uncompressed image yet not exact to the previous one as within the compression process little information related to the image has been lost. So they are normally applied for the photographs. The very natural example of the lossy compression is a JPEG. In lossless compression, it compresses an image by encoding its all information from the actual file, so that if the image gets decompressed again, then it will be the exactly the same as the actual image. For examples, the lossless image compression are PNG and GIF i.e., GIF only provides 8-bit images. At the time of using a specific format of image compression that basically based on what is being get compressed.

- **Lossless Compression :**

In the lossless compression technique, the decompressed image

will be the same replica of the original image. In this case, the binary data like the documents, executable etc. are compressed. This requires to be reproduced exactly when it gets decompressed again. On the contrary, the images and the music also requires not to be generated 'exactly'. A resemblance of the actual image is sufficient for the most objective, as far as the error or problems between the actual and compressed image is avoidable or tolerable. These types of compression are also known as noiseless as they never add noise to signal or image. It is also termed as the entropy coding as it uses the techniques of decomposition/statistics to remove/reduce the redundancy. It is also used only for some specific applications along with the rigid needs like a medical-imaging. Following are the lossless compression techniques.

- Huffman encoding
- Run length encoding
- Arithmetic coding
- Dictionary Techniques
  - LZ77
  - LZ78
  - LZW
- Bit Plane coding

- **Lossy Compression :**

In the technique of lossy compression, it reduces the bits by recognizing the less required information and by eliminating it. The system of decreasing the size of the file or data is commonly termed as the data-compression, though its formal name is the source-coding, that is, coding gets done at source of data before it gets stored or sent. In these methods some loss of the information is acceptable. Dropping non-essential information from the source of data can save the storage area. The Lossy data-compression methods are known the researches as to on how the people anticipate in the question. As an example, the human eye is very sensitive to the slight variations in the luminance as although there

are so many variations in the color. The Lossy image compression technique is used in the digital cameras, to raise the storage ability with the minimal decline of the quality of the picture. Similarly in the DVDs use the lossy MPEG-2 Video codec technique for the compression of the video. In the lossy audio compression, the techniques of psycho acoustics have been used to eliminate the non-audible or less audible components of signal.



CHECK YOUR PROGRESS

Q 1: Why is image compression required?

.....

.....

Q 2: Why is data redundancy necessary?

.....

.....

Q 3: What are the categories of data redundancy?

.....

.....

Q 4: What are the different types of data compression scheme?

.....

.....

Q 5: What is lossless compression?

.....

.....

Q 6: What is lossy compression?

.....

.....

12.6 HUFFMAN CODING

The most popular technique for removing coding redundancy is due to Huffman (Huffman [1952]). When coding the symbols of an information source individually, Huffman coding yields the smallest possible number of code symbols per source symbol. In terms of the noiseless

coding theorem, the resulting code is optimal for a fixed value of n, subject to the constraint that the source symbols be coded one at a time.

The first step in Huffman's approach is to create a series of source reductions by ordering the probabilities of the symbols under consideration and combining the lowest probability symbols into a single symbol that replaces them in the next source reduction. Figure 12.3 illustrates this process for binary coding (K-array Huffman codes can also be constructed). At the far left, a hypothetical set of source symbols and their probabilities are ordered from top to bottom in terms of decreasing probability values. To form the first source reduction, the bottom two probabilities, 0.06 and 0.04, are combined to form a "compound symbol" with probability 0.1. This compound symbol and its associated probability are placed in the first source reduction column so that the probabilities of the reduced source are also ordered from the most to the least probable. This process is then repeated until a reduced source with two symbols (at the far right) is reached.

The second step in Huffman's procedure is to code each reduced source, starting with the smallest source and working back to the original source. The minimal length binary code for a two-symbol source, of course, is the symbols 0 and 1. As Fig. 12.3 shows, these symbols are assigned to the two symbols on the right (the assignment is arbitrary; reversing the order of the 0 and 1 would work just as well). As the reduced source symbol with probability 0.6 was generated by combining two symbols in the reduced source to its left, the 0 used to code it is now assigned to both of these symbols, and a 0 and 1 are arbitrarily appended to each to distinguish them from each other. This operation is then repeated for each reduced source until the original source is reached. The final code appears at the far left in Fig. 12.3. The average length of this code is

$$L_{avg} = (0.4)(1)+(0.3)(2)+(0.1)(4)+(0.06)(5)+(0.04)(5)$$
$$=2.2 \text{ bits/symbol} \quad \ldots\ldots(12.5)$$

and the entropy of the source is 2.14 bits/symbol. The resulting Huffman code efficiency is 0.973.

Figure 12.2: Huffman source reductions

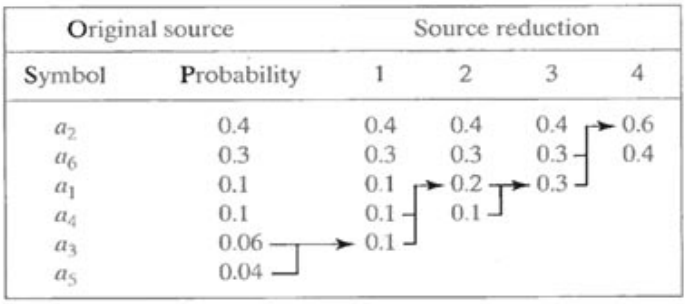
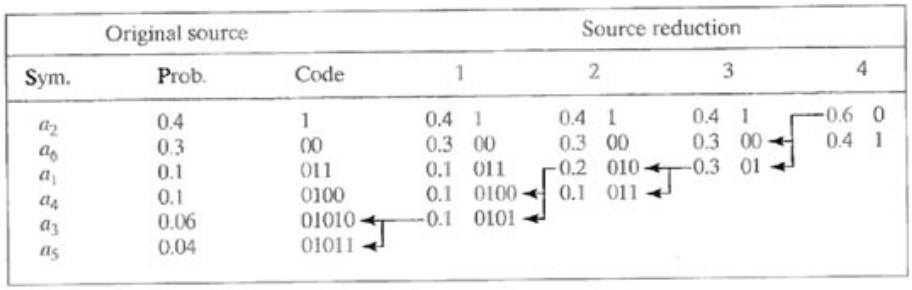


Figure 12.3: Huffman code assignment procedure.



Huffman's procedure creates the optimal code for a set of symbols and probabilities subject to the constraint that the symbols be coded one at a time. After the code has been created, coding and/or decoding is accomplished in a simple lookup table manner. The code itself is an instantaneous uniquely decodable block code. It is called a block code because each source symbol is mapped into a fixed sequence of code symbols. It is instantaneous, because each code word in a string of code symbols can be decoded without referencing succeeding symbols. It is uniquely decodable, because any string of code symbols can be decoded in only one way. Thus, any string of Huffman encoded symbols can be decoded by examining the individual symbols of the string in a left to right manner. For the binary code of Fig. 12.3, a left-to-right scan of the encoded string 010100111100 reveals that the first valid code word is 01010, which is the code for symbol a_3 .The next valid code, is 011, which corresponds to symbol a_1. Continuing in this manner reveals the completely decoded message to be $a_3a_1a_2a_2a_6$.

12.7 TRANSFORM BASE CODING

Transform coding is a technique that is based on modifying the transform of an image. In this technique, a reversible, linear transform is used to map the image into a set of transform coefficients, which are then quantized and coded. For most natural images, a significant number of the coefficients have small magnitudes and can be coarsely quantized with little image distortion. A variety of transformations, including the discrete Fourier transform can be used to transform the image data.

Figure 12.4: A transform coding system: (a) encoder; (b) decoder.

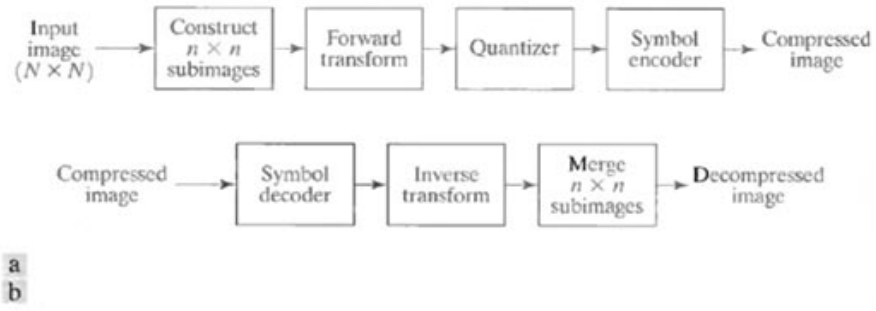


Figure 12.4 shows a typical transform coding system. The decoder implements the inverse sequence of steps (with the exception of the quantization function) of the encoder, which performs four relatively straightforward operations: subimage decomposition, transformation, quantization, and coding. An N X N input image first is subdivided into subimages of size n X n, which are then transformed to generate $(N/n)^2$ subimage transform arrays, each of size n X n. The goal of the transformation process is to decorrelate the pixels of each subimage, or to pack as much information as possible into the smallest number of transform coefficients. The quantization stage then selectively eliminates or more coarsely quantizes the coefficients that carry the least information. These coefficients have the smallest impact on reconstructed subimage quality. The encoding process terminates by coding (normally using a variable-length code) the quantized coefficients. Any or all of the transform encoding steps can be adapted to local image content, called adaptive transform coding, or fixed for all subimages, called nonadaptive transform coding.

12.8 VECTOR QUANTIZATION

Vector quantization is a lossy compression technique used in speech and image coding. In scalar quantization, a scalar value is selected from a finite list of possible values to represent a sample. In vector quantization, a vector is selected from a finite list of possible vectors to represent an input vector of samples. It is a technique that quantizes block of data instead of single sample. The key operation in a vector quantization is the quantization of a random vector by encoding it as a binary code word. Each input vector can be viewed as a point in an *n*-dimensional space. The vector quantizer is defined by a partition of this space into a set of non overlapping *n*-dimensional regions. The vector is encoded by comparing it with a code book consisting of a set of stored reference vectors known as ***code vectors***. The optimality criterion is that a quantization region should consist of all vectors that are closer to its code vector than any of the other code vectors, and the code vector should be the average of all vectors that are in the quantization region.

Let *k* be the number of quantization regions in the *n*-dimensional space. In other words, since every region has its own code vector, *k* is the number of code vectors in the codebook, and *n* represents the number of samples in a code vector. For the number of bits per sample *R*, we therefore have the following:

$$R = \log_2 kn \quad \text{.....(12.6)}$$

Figure 12.5. Vector quantization: two-dimensional decision regions.

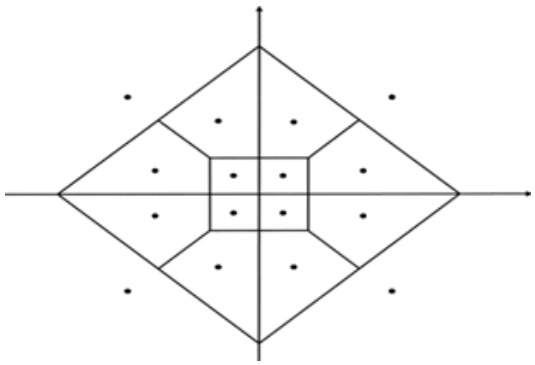


Figure12.5 shows an example of a vector quantization with k=16 and n=2 (i.e., R=2 bits per sample). Assuming that the size of

the codebook is sufficiently large, it can be shown that the SNR_o for the vector quantizer is given by:

$$SNR_o = 6R+C_n dB \ldots\ldots\ldots\ldots\ldots\ldots\ldots (12.7)$$

where C_n is a constant (expressed in dB) that increases with the dimension *n*. Note that the SNR_o once again follows the 6-dB law (i.e., it increases approximately 6 dB when there is an increase of one bit in the number of bits per sample). However, the constant term C_n has a higher value than the constant used in scalar quantization. This additional improvement is due to the fact that the vector quantization optimally exploits the correlations among the samples in a vector, but at the expense of increased complexity. Due to the complexity of the codebook search, vector quantization is not quite prevalent.



CHECK YOUR PROGRESS

Q 7: What is output we get from Huffman coding?

.....

.....

Q 8: What are the intermediate processes of compressed image in transform coding?

.....

.....

Q 9: What is the basic principle of transform coding?

.....

.....

Q 10: What is the key operation of vector quantization?

.....

.....



12.9 LET US SUM UP

- Compression is compact representation.
- Image compression deals with the representation of an image with minimum number of bits.

- Compression is achieved by minimizing the redundancies in an image.
- Different types of redundancies in an image are i) coding redundancy ii) interpixel redundancy iii) psychovisual redundancy.
- Two broad categories of compression are i) lossless compression and ii) lossy compression.
- The techniques commonly used in lossless compression: Huffman encoding, Run length encoding, Arithmetic coding, Dictionary Technique like LZW and Bit Plane coding
- Transform coding is based on modifying the transform of an image. A reversible, linear transform is used to map the image into a set of transform coefficients.
- Quantization is basically approximation. Quantization is a non-linear, irreversible process and classified as scalar quantization and vector quantization.



12.10 FURTHER READING

1. Gonzales, R. C., & Woods, R. E. (2008). Digital image processing. Prentice Hall.
2. Jayaraman, S. Esakkirajandan, S., & Veerakumar, T. (2009). Digital Image Processing, New Delhi: Tata McGraw-Hill Education Private Limited.

12.11 ANSWERS TO THE CHECK YOUR PROGRESS

Ans to Q No 1: In a distributed environment large image files remain a major bottleneck within systems. Compression is an important component of the solutions available for creating file sizes of manageable and transmittable dimensions.

Ans to Q No 2: Data redundancy is a main issue in digital image compression. It is not an abstract concept but a mathematically quantifiable entity.

Ans to Q No 3: In digital image compression, three basic data redundancies can be identified and exploited: coding redundancy, interpixel redundancy, and psychovisual redundancy.

Ans to Q No 4: Lossy Compression and Lossless Compression

Ans to Q No 5: In the technique of Lossless compression with the compressing of data that is when get decompressed, will be the same replica of actual data. It is applicable for document type data.

Ans to Q No 6: In the technique of Lossy compression, it decreases the bits by recognizing the not required information and by eliminating it. The system of decreasing the size of the file of data is commonly termed as the data-compression

Ans to Q No 7: Huffman's procedure creates the optimal code for a set of symbols and probabilities subject to the constraint that the symbols be coded one at a time.

Ans to Q No 8: Input imageàConstruct subimageàForward Transformà Quantizer à Encoder àCompressed Image.

Ans to Q No 9: Transform coding is based on modifying the transform of an image.

Ans to Q No 10: The key operation in a vector quantization is the quantization of a random vector by encoding it as a binary code word.



12.12 MODEL QUESTIONS

Q 1: Define image compression.

Q 2: Why image compression is necessary?

Q 3: Explain about the redundancies in a digital image.

Q 4: Explain about image compression scheme.

Q 5: Write about the applications of lossy and lossless compression

Q 6: Explain the method of Huffman codes .

Q 7: Explain with a block diagram about transform coding system.

Q 8: Explain about vector quantization.

Q 9: Calculate the average length of Huffman code for the following symbols whose probability of occurrence is given below:

Symbols	Probability
A_1	0.9
A_2	0.06
A_3	0.02
A_4	0.02

*** ***** ***

UNIT 13 : BINARY IMAGE PROCESSING

UNIT STRUCTURE

13.1 Learning Objectives

13.2 Introduction

13.3 Binarisation

13.4 Mathematical Morphology

13.5 Structuring Element

13.6 Morphological Image Processing

13.7 Basic Set Theory

13.8 Standard Binary Morphological operations

- 13.8.1 Dialation
- 13.8.2 Erosion
- 13.8.3 Opening
- 13.8.4 Closing

13.9 Let Us Sum Up

13.10 Further Reading

13.11 Answers to Check Your Progress

13.12 Model Questions

13.1 LEARNING OBJECTIVES

After going through this unit you will be able to:

- learn about binary image
- describe mathematical morphology and structuring element
- learn how the dialation and erosion work
- describe properties of mathematical morphology

13.2 INTRODUCTION

In the previous unit we have learnt about the image compression which is a process of reducing object size. The different types and techniques of image compression have been discussed in that unit.

In this unit, we will introduce binary image processing. Here we

will find how operations can be performed on binary image. The word ‘binary’ means 0 and 1. So, binary image indicates that the image with pixel values of either 0 or 1.

13.3 BINARISATION

There are basically three types of images i.e. binary image, gray scale image and colour image. Each pixel of binary image contains only one bit(0 or 1); gray scale image contains 8 bits(0 to 255 pixel intensity values) and colour image contains 24 bits(8 bits for each of three colour R G B). Colour image or gray scale image can be converted to binary image. The process of converting an image(colour or gray scale) into binary image is called **binarisation**. There are different techniques used for converting an image into binary. The common technique is global thresholding. Global thresholding sets all pixels above a defined value to white and the rest of the pixels to black in the image. Figure 13.1 shows the example of gray scale and binary images. Figure 13.1(b) has been created using Figure 13.1(a) taking global threasholding values 120.

Figure 13.1: Image examples



a. Gray scale image b. Binary image

13.4 MATHEMATICAL MORPHOLOGY

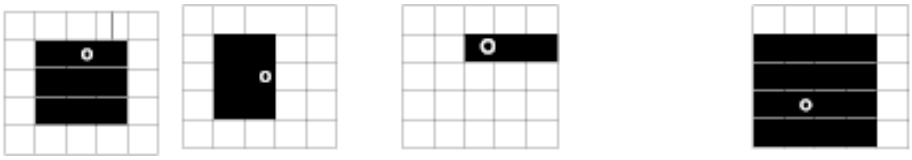
In digital image processing morphology means structure or shape. Mathematical morphology is a tool for finding image components that are useful for representation and description. The content of mathematical morphology is based on set theory. There are many morphological operators used in binary image processing. Morphological operators take a binary image and a mask known as a structuring element as inputs.

Then the set operators such as insersion, union, inclusion and complement can be applied to the images.

13.5 STRUCTURING ELEMENT

Structuring elements are small portion of an image that are used to process the original image. The structuring element can be a small binary image or a flat file that consists of coordinate values. It plays an important role in morphological operations with its different size. Generally the size of structuring element can be 2x2, 3x3, 4x4, 5x5 etc. Figure 13.2 shows the structuring elements of different size. The circle shown in the figure is called the centre pixel, where the calculated value is applied.

Figure 13.2: Structuring elements of different size

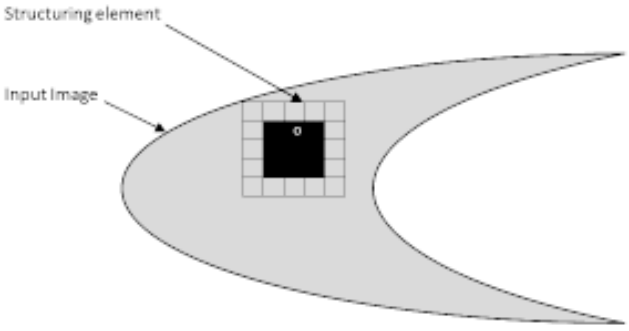


(a) 3x3 (b) 3x2 (c) 1x3 (d) 4x4

13.6 MORPHOLOGICAL IMAGE PROCESSING

Morphological image processing is performed taking the input binary image and the structuring element. The structuring element is moved over the binary image pixel by pixel from left to right and from top to bottom. When the structuring element is centered over a portion of the image, a logical operation is performed on the pixels covered by the structuring element. Figure 13.3 shows the basic idea of morphological image processing using input binary image and structuring element.

Figure 13.3: Morphological image processing



13.7 BASIC SET THEORY

Mathematical morphology is based on basic set theory. This section describes the basic set theory used in image processing.

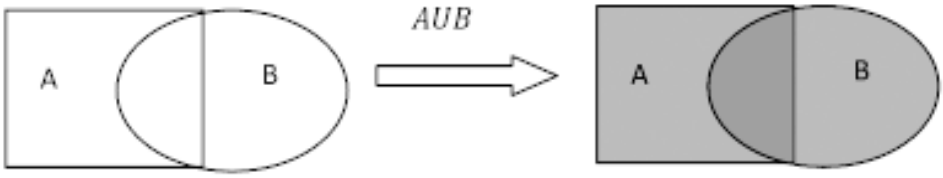
13.7.1 Union

The union of two binary images A and B is written as AUB. AUB is the set whose elements are either the elements of A or B or both. Mathematically, it is written as

$$\mathrm{AUB} = \mathrm{def}\left\{x \middle| x \in A \; or \; x \in B\right\}$$

The union of images A and B has been shown in Fig. 13.4

Figure 13.4: Union of two images A and B

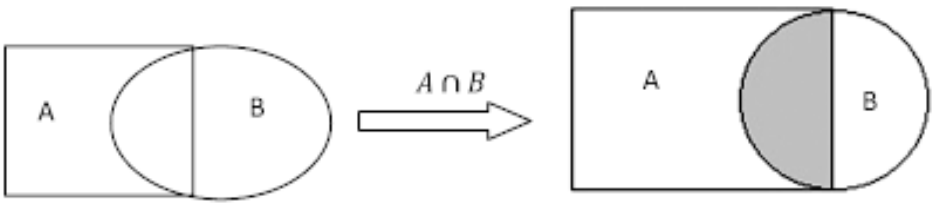


13.7.2 Intersection

The intersection of two binary images A and B is written as A)"B. A)"B is the set whose elements are common to both images A and B. Mathematically, it is written as

The intersection of images A and B has been shown in Fig. 13.5

Figure 13.5: Intersection of two images A and B



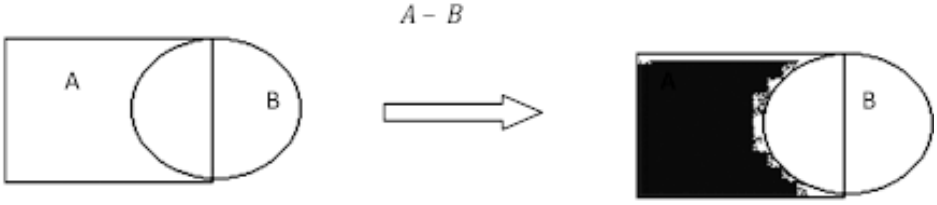
13.7.3 Difference

The difference of two binary images A and B is written as A – B. A – B is the set whose elements are in image A and not

in image B. Mathematically, it is written as

The difference of images A and B has been shown in Fig. 13.6

Figure 13.6: Difference of two images A and B



13.7.4 Complement

The complement of image A is the set consisting of everything but not in image. It is written as A^c. The definition of complement operation is represented by

$A_1^c = \{x \mid x \notin A \text{ and } x \in B\}$

The result of complement operation is shown in figure 13.7

Figure 13.7: Complement of two images A and B



13.7.5 Reflection

Reflection of an image A can be viewed as mirror of that image. It is defined as

The result of reflection operation has been depicted in Fig. 13.8

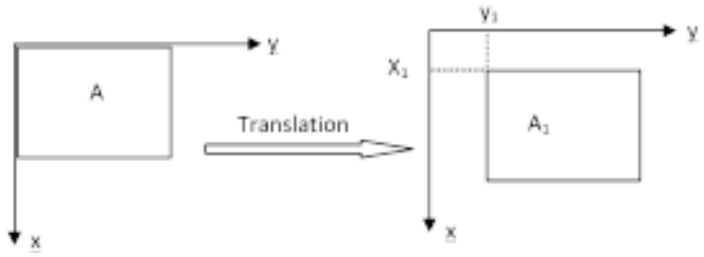
Figure 13.8: Reflection of image A



13.7.6 Translation

The translation operation shifts the origin of image A from point (x,y) to some other point (x_1, y_1). Translation operation of image A is defined as

Figure 13.9: Translation of image A



The result of translation operation has been depicted in Fig. 13.9

13.8 STANDARD BINARY MORPHOLOGICAL OPERATIONS

The basic morphological operations are Dialation and Erosion.

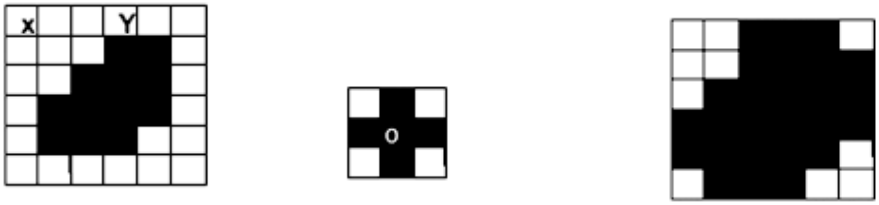
13.8.1 Dialation

Dilation is one of the two basic operators in the area of mathematical morphology. It is typically applied to binary and gray scale images. The basic effect of this operator on a binary image is that it gradually increases the boundaries of regions of the image. The way the image increases is determined by the structuring element. The structuring element is small in size normally 3x3, 5x5, 7x7 etc. The dialation operation is performed using the structuring element shifted from left to right and from top to bottom. At each shift, the process will look for any overlapping similar pixels between the structuring element and the binary image. If there exists any overlapping the pixel under the centre position of the structuring element will be turned to 1 or black. Let us define X as the reference image and B as the structuring element. The dialation operation is defined by

$$X \oplus B = \{ z \mid (\hat{B})_z \cap X \subseteq X \} \quad (13.1)$$

Where $\hat{B}$ is the image B rotated about the origin. Equation 13.1 states that when the image X is dialated by the structuring element B, the outcome element z would be that there will be atleast one element in B that intersects with an element in X. For example,

Figure 13.10: Dialation operation using 3x3 structuring element



(a) Input image (b) Structuring element (c) Dialated image

Figure 13.10 shows the dialation operation. The black square represent 1 and the white square represent 0. Initially, the centre of the structuring element is placed at position X in Figure 13.10(a). At this point, there is no overlapping between any black square of B and the black square of X. So, at position X the square will remain white. The structuring element will then be shifted by one pixel right. The shifting process will continue till the last pixel of the row. At position Y, one black square(pixel) of B is overlapping or intersecting with the black square of X. Thus, at position Y the square will be changed to black. The structuring element B is shifted from left to right and from top to bottom on the image X to get the dialated image as shown in Figure 13.10(c).

Program 13.1: Python code for image dialation

```
# coding: utf-8
# Image dialation using 3x3, 5x5 and 7x7
structuring element

import cv2
import numpy as np
from matplotlib import pyplot as plt
```

```
# Reading the image
img=cv2.imread('/home/rusa/Desktop/j.png',0)
#Declaring the structuring element
kernel1 = np.ones((3,3),np.uint8)
kernel2 = np.ones((5,5),np.uint8)
kernel3 = np.ones((7,7),np.uint8)

# Performing dialation operation
dilation1 = cv2.dilate(img, kernel1, iterations
= 1)
dilation2 = cv2.dilate(img, kernel2, iterations
= 1)
dilation3 = cv2.dilate(img, kernel3, iterations
= 1)

# Display the images
plt.subplot(221), plt.imshow(img)
plt.axis("off")
plt.subplot(222), plt.imshow(dilation1)
plt.axis("off")
plt.subplot(223), plt.imshow(dilation2)
plt.axis("off")
plt.subplot(224), plt.imshow(dilation3)
plt.axis("off")
```

Output :



Original image



Dialation using 3x3 structuring element





Dialation using 5x5 and 7x7 structuring element

13.8.2 Erosion

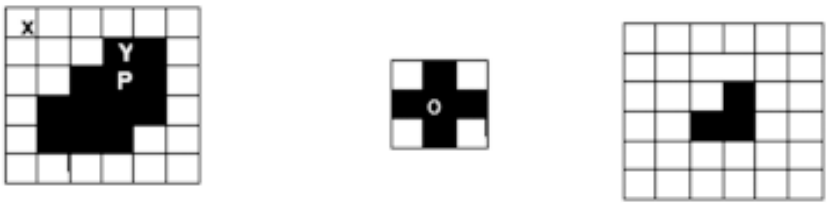
Erosion is the reverse process of dialation. Dilation increases the boundary region of an image whereas erosion shrinks the image. The amount of reduction depends upon the size of the structuring element. The structuring element is normally smaller than the image. Smaller the structuring element lesser the computation time. Similar to the dialation process, the erosion process shifts the structuring element from left to right and top to bottom. At the centre position, indicated by the centre of the structuring element, the process will look for whether there is a complete overlap with the structuring element or not. If there is no complete overlapping then the centre pixel indicated by the centre of the structuring element will be set white or 0.

Mathematically, it may be defined as

$$X \ominus B = \{z \mid (B)_Z \subseteq X\} \qquad (13.2)$$

Where X is reference binary image, B is structuring element, Equation 13.2 states that the outcome element z is considered only when the structuring element is a subset or equal to the binary image X. For example,

Figure 13.11: Erotion operation using 3x3 structuring element



(a) Input image (b) Structuring element (c) Eroded image

Figure 13.11 shows the erosion operation. The black square represents 1 and the white square represent 0. Initially, the centre of the structuring element is placed at position X in Figure 13.11(a). At this point, there is no overlapping between any black square of B and the black square of X. So, at position X, the square will remain white. The structuring element will then be shifted by one pixel right. The shifting process will continue till

the last pixel of the row. At position Y, complete overlapping or intersecting with the black square of X is not found, thus the black square marked with Y will be turned to white. The structuring element is then shifted further until its centre reaches the position P. At position P overlapping is complete, that is, all the black squares in the structuring element overlap with the black squares in the image. Thus, the centre of the structuring element corresponding to the image will be black. The structuring element B is shifted from left to right and from top to bottom on the image X to get the eroded image as shown in Figure 13.11(c).

Program 13.2: Python code for image erosion

```
# coding: utf-8

# Image erosion using 3x3, 5x5 and 7x7 structuring element

import cv2
import numpy as np
from matplotlib import pyplot as plt
#Reading the image from a folder in binary mode
img=cv2.imread('/home/rusa/Desktop/j.png',0)
#defining the structuring element
kernel1 = np.ones((3,3),np.uint8)
kernel2 = np.ones((5,5),np.uint8)
kernel3 = np.ones((7,7),np.uint8)
# erosion operation
erosion1 = cv2.erode(img, kernel1, iterations = 1)
erosion2 = cv2.erode(img, kernel2, iterations = 1)
erosion3 = cv2.erode(img, kernel3, iterations = 1)

# Display the output in four window
```

```
plt.subplot(221), plt.imshow(img)
plt.axis("off")
plt.subplot(222), plt.imshow(erosion1)
plt.axis("off")
plt.subplot(223), plt.imshow(erosion2)
plt.axis("off")
plt.subplot(224), plt.imshow(erosion3)
plt.axis("off")
```

Output:




Original image Erosion using 5x5 and 7x7 structuring element




Erosion using 3x3 structuring element

13.8.3 Opening

Opening is based on the morphological operations, erosion and dialation. Opening smoothes the inside of the object conture, breaks narrow strips and eliminates thin portions of the image. It is done by first applying erosion and then dialation operations on the image. Mathematically, the opening process can be defined as

$$X \circ B = (X \oplus B) \ominus B \qquad (13.3)$$

13.8.4 Closing

The closing operation is the reverse of opening operation. It is a dialation operation followed by an erosion operation. The closing operation fills the small holes and gaps in a single pixel

object. It has the same effect as the opening operation, in that it smoothes contours and maintains shapes and size of objects. Mathematically, the closing process can be defined as

$$X \circ B = (X \oplus B) \ominus B$$



CHECK YOUR PROGRESS

Q 1: Write True or False

a. Each pixel of a binary image can take only 0 and 1 value

b. A grayscale image pixel can take only one value from 256 colour intensities irrespective of colour code

c. Size of structuring element can be larger than reference input image

d. Morphological operations are used for colour conversion

e. Set operation can be performed on colour image

f. Dialation operation shrinks the original image

g. Structuring element can be of any shape



13.9 LET US SUM UP

- There are basically three types of images, binary, gray scale and colour.
- Binarisation is the process of converting a gray scale image into binary image.
- Mathematical morphology is a tool for finding image components that are useful for representation and description.
- Structuring element is a set of pixels normally in numbers 9, 25, 49 etc.
- There are four basic morphological operations mainly used in image processing i.e. dialation, erosion, opening and closing.
- Dialation operation increases the size of an image
- Erosion operation reduces the size of an image. The increase or shrink of an image depends on the structuring element



13.10 ANSWERS TO CHECK YOUR PROGRESS

Ans to Q No 1:

a. True
b. False
c. False
d. False
e. True
f. False
g. False



13.11 FURTHER READING

1. Gonzales, R. C., & Woods, R. E. (2008). Digital image processing. Prentice Hall.
2. Jayaraman, S. Esakkirajandan, S., & Veerakumar, T. (2009). Digital Image Processing, New Delhi: Tata McGraw-Hill Education Private Limited.



13.12 MODEL QUESTIONS

Q 1: What is binary image? What are the differences between binary and grayscale image?

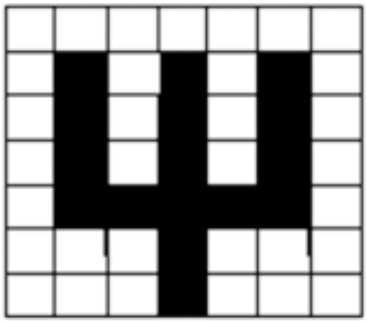
Q 2: What is mathematical morphology? How is the morphology used in image processing?

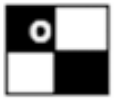
Q 3: Explain the different set theories used in image processing.

Q 4: Explain the different logical operations used in image processing.

Q 5: Explain any one standard binary morphological operations used in image processing.

Q 6: Using the input image and structuring element as given below, fi





Input image Structuring element

*** ***** ***

UNIT 14 : COLOR IMAGE PROCESSING

UNIT STRUCTURE

14.1 Learning Objectives

14.2 Introduction

14.3 Color Formation

14.4 Human Perception of Color

14.5 Color Model

- 14.5.1 RGB Model
- 14.5.2 CMK Model
- 14.5.3 HSI Model
- 14.5.4 YIQ Model

14.6 Histogram Equalization of Color Image

14.7 Color Image Filtering

14.8 Let Us Sum Up

14.9 Further Reading

14.10 Answers to Check Your Progress

14.11 Model Questions

14.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- learn about the fundamentals of colour formation
- describe different colour models
- describe histogram of colour image

14.2 INTRODUCTION

In color image processing, it is motivated by two principles: first is color, which is a powerful description and which simplifies the object identification and extraction from a scene. And second the human perception about the color. Human eye can visualize a range of colors and a shade of gray color. Color image processing is divided into two major areas: full color processing and pseudo color processing. In the

first the images queried with a full color sensor and in the second category, the problem is accessing a color to a particular monochromatic intensity or a range of intensities. Most of image processing is done at the pseudo color level. In the past decades, color processing hardware are become available in reasonable prices. As a result full color processing techniques are now used in wide range of application, including publishing and the internet.

14.3 COLOR FORMATION

In 1666, Sir Isaac Newton discovered that when a beam of sunlight passes through the glass prism, the emerging light is not white but consists of spectrum color ranging from violet at one end to red in the other end. When viewed in full color, no color in the spectrum ends abruptly, but rather each color blends smoothly in the next.

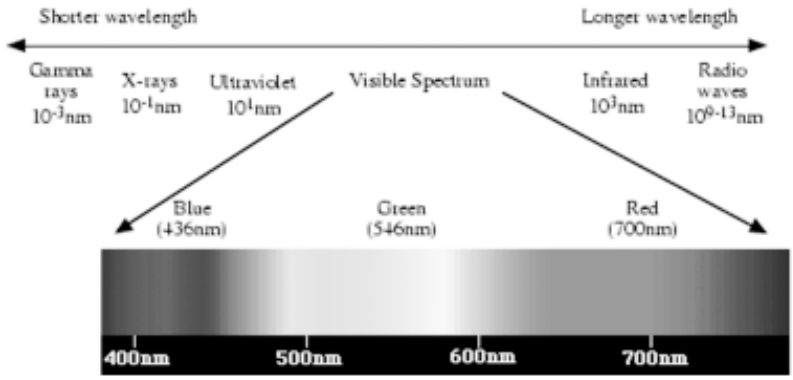
Primarily color depends on the reflectance properties of an object. We see only those rays that are reflected, while others are absorbed. However, we also must consider the color of the light source, and the nature of human visual system. For example, an object that reflects both red and green will appear green when there is green but no red light illuminating it, and conversely it will appear red in the absence of green light. In pure white light, it will appear yellow (= red + green).

14.4 HUMAN PERCEPTION OF COLOR

The human visual system can distinguish hundreds of thousands of different color shades and intensities, but only around 100 shades of grey. Therefore, in an image, a great deal of extra information may be contained in the color, and this extra information can then be used to simplify image analysis, e.g. object identification and extraction based on color.

Three independent quantities are used to describe any particular color. The *hue* is determined by the dominant wavelength. Visible colors occur between about 400nm (violet) and 700nm (red) on the electromagnetic spectrum, as shown in figure 14.1

Figure 14.1: The visible spectrum.



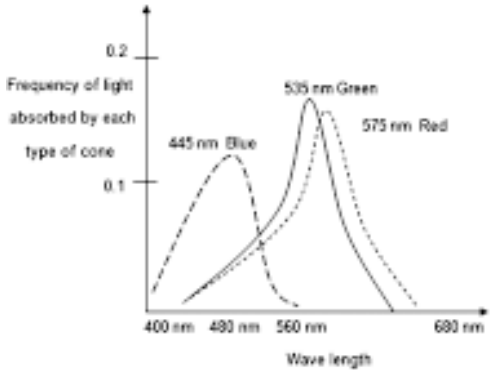
The *saturation* is determined by the excitation purity, and it depends on the amount of white light is mixed with the hue. A pure hue is fully saturated, i.e. no white light mixed in. Hue and saturation together determine the *chromaticity* for a given color. Finally, the *intensity* is determined by the actual amount of light, with more light corresponding to more intense colors.

Achromatic light has no color - its only attribute is quantity or intensity. Greylevel is a measure of intensity. The *intensity* is determined by the energy, and is, therefore a physical quantity. On the other hand, *brightness* or *luminance* is determined by the perception of the color, and is, therefore, psychological. Given equally intense blue and green, the blue is perceived as much darker than the green.

- **Tristimulus Theory of Color Perception**

  The human retina has 3 kinds of cones. The response of each type of cone as a function of the wavelength of the incident light is shown in figure 14.2. The peaks for each curve are at 440nm (blue), 545nm (green) and 580nm (red) as shown in Fig.14.2. The last two actually peak in the yellow part of the spectrum.

Figure 14. 2: Spectral response curves for each cone type

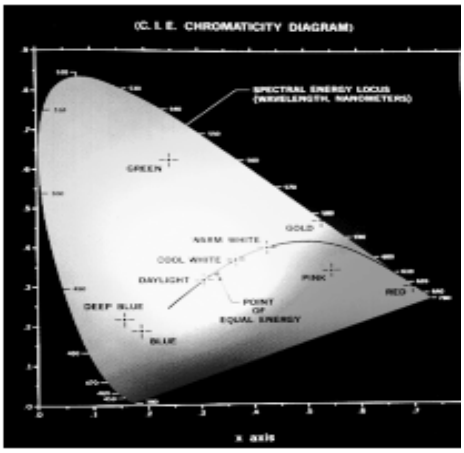


- **CIE Primaries**

The tristimulus theory of color perception seems to imply that any color can be formed from a mix of the three primary colors as red, green and blue, but some colors can not be matched in this way. If one of the primaries is added to one of these unmatchable colors, it can be matched by a mixture of the other two, and so the color may be considered to have a negative weighting of that particular primary.

In 1931, the *Commission Internationale de l'Éclairage (CIE)* defined three standard primaries, called *X*, *Y* and *Z* that can be added to form all visible colors. The primary *Y* was chosen so that its color matching function exactly matches the luminous-efficiency function for the human eye, given by the sum of the three curves in figure 14.3. Here x and y are the normalized amounts of the X and Y primaries present, and hence z = 1 - x - y gives the amount of the Z primary required.

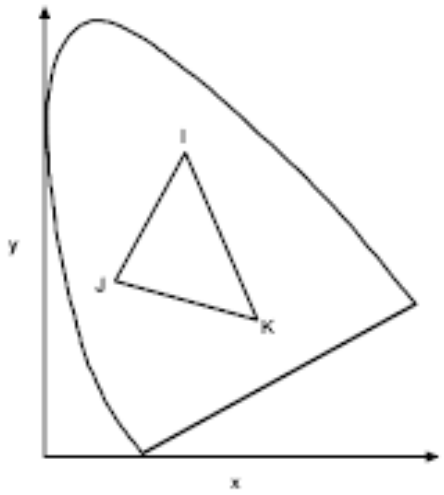
Figure 14.3: CIE Chromaticity Diagram



The CIE Chromaticity Diagram (Fig. 14.3) shows all visible colors. The x and y axis give the normalized amounts of the X and Y primaries for a particular color, and hence z = 1 - x + y gives the amount of the Z primary required. Chromaticity depends on dominant wavelength and saturation, and is independent of luminous energy. Colors with the same chromaticity, but different luminance all map to the same point within this region.

The pure colors of the spectrum lie on the curved part of the boundary, and a standard white light has color defined to be near (but not at) the point of equal energy x = y = z = $^1/_3$. Complementary colors, i.e. colors that add to give white, lie on the endpoints of a line through this point. As illustrated in figure 14.4, all the colors along any line in the chromaticity diagram may be obtained by mixing the colors on the end points of the line. Furthermore, all colors within a triangle may be formed by mixing the colors at the vertices. This property illustrates graphically the fact that all visible colors cannot be obtained by a mix of R, G and B (or any other three visible) primaries alone, since the diagram is not triangular!

Figure 14.4: Mixing colors on the chromaticity diagram. All colors on the line IJ can be obtained by mixing colors I and J, and all colors in the triangle IJK can be obtained by mixing colors I, J and K.



14.5 COLOR MODEL

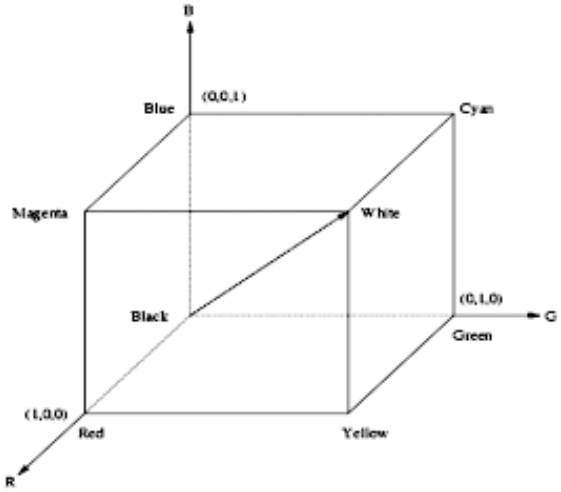
Color models provide a standard way to specify a particular color, by defining a 3D coordinate system and a subspace that contains all

constructible colors within a particular model. Any color that can be specified using a model will correspond to a single point within the subspace it defines. Each color model is oriented towards either specific hardware (RGB,CMY,YIQ), or image processing applications (HSI). There are many models available as the colour science is a broad field that covers many areas of application.

14.5.1 The RGB Model

In the RGB model, an image consists of three independent image planes, one in each of the primary colors: red, green and blue. Specifying a particular color is by specifying the amount of each of the primary components present. Figure 14.5 shows the geometry of the RGB color model for specifying colors using a Cartesian coordinate system. The grayscale spectrum, i.e., those colors made from equal amounts of each primary, lies on the line joining the black and white vertices.

Figure 14.5: The RGB color cube. The grayscale spectrum lies on the line joining the black and white vertices.



This is an *additive* model, i.e. the colors present in the light add to form new colors, and is appropriate for the mixing of colored light for example. The image on the left of figure 14.5 shows the additive mixing of red, green and blue primaries to form the three secondary colors yellow (red + green), cyan (blue + green) and

magenta (red + blue), and white (red + green + blue). The RGB model is used for color monitors and most video cameras.

14.5.2 The CMY Model

The CMY (cyan-magenta-yellow) model is a *subtractive* model appropriate to absorption of colors, for example due to pigments in paints. Whereas the RGB model asks what is added to black to get a particular color, 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 colors (see the image on the right of figure 14.6).

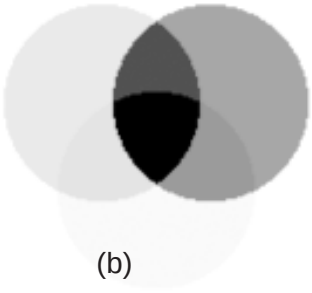
When a surface coated with cyan pigment is illuminated by white light, no red light is reflected, and similarly for magenta and green, and yellow and blue. The relationship between the RGB and CMY models is given by:

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

The CMY model is used by printing devices and filters.

Figure14.6(a) shows the additive mixing of red, green and blue primaries to form the three secondary colors yellow (red + green), cyan (blue + green) and magenta (red + blue), and white (red + green + blue). Figure 14.6(b) shows the three subtractive primaries, and their pairwise combinations to form red, green and blue, and finally black by subtracting all three primaries from white.

Figure 14.6: Primary and Secondary colours of light

(a)  (b) 

- **Why does blue paint plus yellow paint give green?**

  As all school children know, the way to make green paint is to mix blue paint with yellow. But how does this work? If blue paint absorbs all but blue light, and yellow absorbs blue only, when combined no light should be reflected and black paint should result in.

  However, what actually happens is that imperfections in the paint are exploited. In practice, blue paint reflects not only blue, but also some green. Since the yellow paint also reflects green (since yellow = green + red), some green is reflected by both pigments, and all other colors are abosrbed, resulting in green paint.



CHECK YOUR PROGRESS

Q 1: What is the range of human visible color in electromagnetic spectrum

.....

.....

Q 2: What are the basic property(s) define color?

.....

.....

Q 3: What is Tristimulus Theory?

.....

.....

Q 4: What is CIE primaries?

.....

.....

Q 5: What is RGB color model?

.....

.....

Q 6: What is CMY color model?

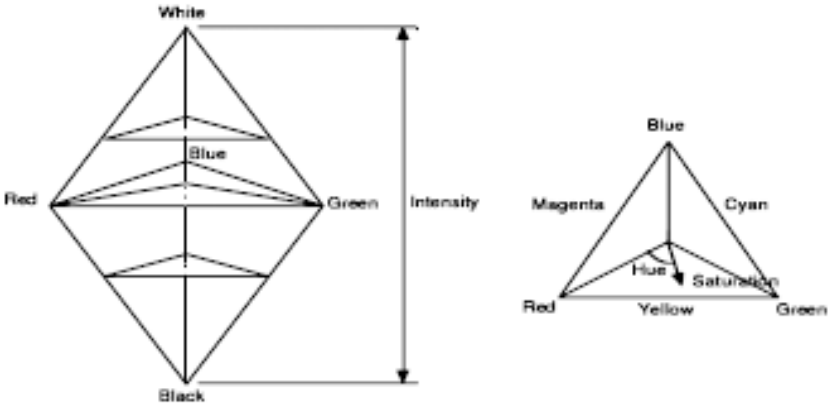
.....

.....

14.5.3 The HSI Model

We know that color is specified by the three quantities hue, saturation and intensity. This is the HSI model, and the entire space of colors that may be specified in this way as shown in figure 14.7. Figure 14.7 shows the HSI solid on the left, and the HSI triangle on the right, formed by taking a horizontal slice through the HSI solid at a particular intensity. Hue is measured from red, and saturation is given by distance from the axis. Colors on the surface of the solid are fully saturated, i.e. pure colors, and the greyscale spectrum is on the axis of the solid. For these colors, hue is undefined.

Figure 14.7: HIS model



Conversion between the RGB model and the HSI model is quite complicated. The intensity is given by

$$I = \frac{R+G+B}{3}$$

where the quantities R, G and B are the amounts of the red, green and blue components, normalized to the range [0,1]. The intensity is therefore just the average of the red, green and blue components. The saturation is given by:

$$s = 1 - \frac{\min(R,G,B)}{1} = 1 - \frac{3}{R+G+B} MIN(R,G,B)$$

where the min(R,G,B) term is really just indicating the amount of white present. If any of R, G or B are zero, there is no white and we have a pure color.

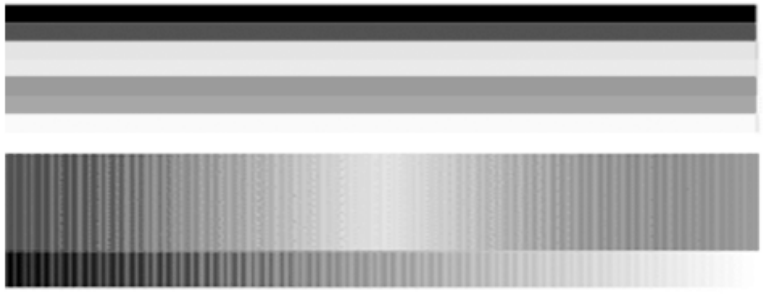
14.5.4 The YIQ Model

The YIQ (luminance-inphase-quadrature) model is a recoding of RGB for color television, and is a very important model for color image processing. The conversion from RGB to YIQ is given by:

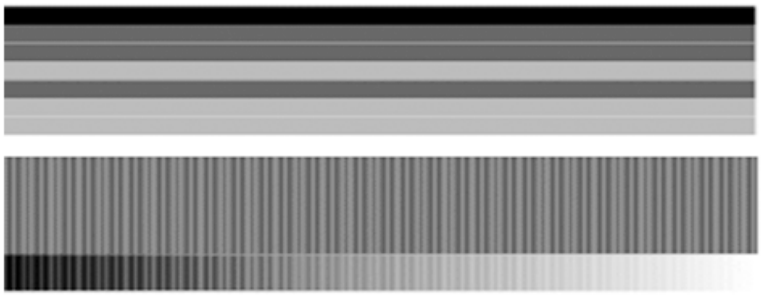
$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The luminance (Y) component contains all the information required for black and white television, and captures our perception of the relative brightness of particular colors. That we perceive green as much lighter than red, and red lighter than blue, is indicated by their respective weights of 0.587, 0.299 and 0.114 in the first row of the conversion matrix above. These weights should be used when converting a color image to greyscale if you want the perception of brightness to remain the same. This is not the case for the intensity component in an HSI image, as shown in figure 14.8.

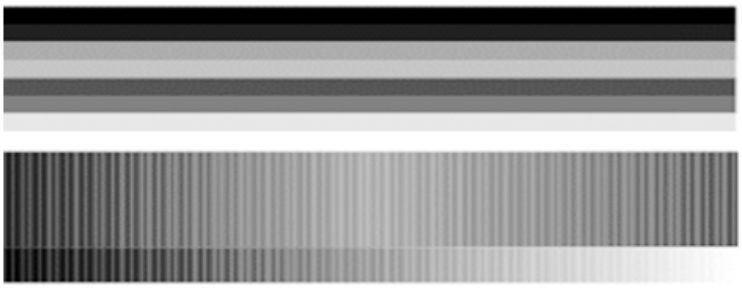
Figure 14.8: Image (a) shows a color test pattern, consisting of horizontal stripes of black, blue, green, cyan, red, magenta and yellow, a color ramp with constant intensity, maximal saturation, and hue changing linearly from red through green to blue, and a greyscale ramp from black to white. Image (b) shows the intensity for image (a). Note how much detail is lost. Image (c) shows the luminance. This third image accurately reflects the brightness variations preceived in the original image.



(a) Colour Image



(b) Intensity Image



(c) Luminance Image

- **Applying Grayscale Transformations to Color Images:** Given all these different representations of color, and hence color images, the question arises as to what is the best way to apply the image processing techniques we can apply on these images? One possibility is to apply the transformations to each color plane in an RGB image, but what exactly does it mean? If we want to increase the contrast in a dark image by histogram equalization, can we just equalize each color independently? This will result in quite different colors in our transformed image. In general, it is better to apply the transformation just to the intensity component of an HSI image, or the luminance component of a YIQ image, thus leaving the chromaticity unaltered.

  An example is shown in figure 14.9. When histogram equalization is applied to each color plane of the RGB image, the final image is lighter, but also quite differently colored to the original. When histogram equalization is only applied to the luminance component of the image in YIQ format, the result is more like a lighter version of the original image, as required.

Figure 14.9



In figure 14.9, the top image is a very dark image of a forest scene. The middle image is the result of applying histogram equalization to each of the red, green and blue components of the original image. The bottom image is the result of converting the image to YIQ format, and applying histogram equalization to the luminance component only.

The eye perceives color when certain wavelengths of light are reflected off the object and enter the eye through the lens. The message is then sent to the retina where photoreceptors in the shape of cones or rods interpret the message. Color vision is interpreted by the cones.

Within the cones are three kinds of photosensitive receptors that interpret different wavelengths based on the makeup. Normally, each cone is responsible for interpreting a different range of wavelengths: long, medium or small, or red,

green, or blue sensitive wavelengths. The LMS title is more accurate because the capabilities of each cone overlap. While the cones responsibilities do overlap, a cone not interpreting its range of responsibly could lead to an inability to perceive or distinguish certain colors.

14.6 Histogram Equalization

Histogram of an image, like other histograms, also shows frequency. In an image histogram, the x axis shows the gray level intensities and the y axis shows the frequency of these intensities.

- **Applications of Histograms**
  - Histograms have many uses in image processing. The first use is the analysis of the image. We can predict about an image by just looking at its histogram.
  - The second use of histogram is for brightness purposes. The histograms have wide application in image brightness. Not only in brightness, but histograms are also used in adjusting contrast of an image.

Another important use of histogram is to equalize an image.

- PMF stands for probability mass function. It gives the probability of each number in the dataset or you can say that it basically gives the count or frequency of each element.
- CDF stands for cumulative distributive function. It is a function that calculates the cumulative sum of all the values that are calculated by PMF. It basically sums the previous one.
- Histogram equalization is used for enhancing the contrast of the images.

PMF and CDF are both used in histogram equalization as the first and the second step. Since in histogram equalization, we have to equalize all the pixel values of an image, so PMF helps us in calculating the probability of each pixel value in an image. And CDF gives us the cumulative sum of these values. Further on, this CDF is multiplied by levels, to find the new pixel intensities, which

are mapped into old values, and thus the histogram is equalized.

Histogram equalization is used to enhance contrast. It is not necessary that contrast will always be increased in this. There may be some cases were histogram equalization can be worse. In that case the contrast is decreased.

Now, we will perform histogram equalization to it.

The method is useful in images with backgrounds and foregrounds that are both bright or both dark. In particular, the method can lead to better views of bone structure in x-ray images, and to better detail in photographs that are over or under-exposed. A key advantage of the method is that it is a fairly straightforward technique and an invertible operator. So, in theory, if the histogram equalization function is known, then the original histogram can be recovered. The calculation is not computationally intensive. A disadvantage of the method is that it is indiscriminate. It may increase the contrast of background noise, while decreasing the usable signal.

Histogram equalization often produces unrealistic effects in photographs; however it is very useful for scientific images like thermal, satellite or x-ray images, often the same class of images to which one would apply false-color. Also histogram equalization can produce undesirable effects applied to images with low color depth. Histogram equalization will work best when applied to images with much higher color depth than palette size, like continuous data or 16-bit gray-scale images.

Histogram Equalization can be considered as redistribution of the intensity of the image. Color histogram equalization can be achieved by converting a color image into HSV/HSI image and enhancing the Intensity while preserving hue and saturation components.

However, performing histogram equalization on components of R,G and B independently will not enhance the image. This section describes how to implement the histogram equalization method in Python.

OpenCV provides us a function through which we can apply histogram equalization on an image, namely equalizeHist(). It is applicable to a grayscale image but for a image with three channels (RGB), we cannot apply histogram equalization on the three channels in a separate manner. As a solution to this problem, we can convert our image to the YUV color space, equalize theY channel, and finally convert the result to RGB. We use BGR instead of RGB here, since OpenCV loads the images in BGR format. equalization on pout.jpg looks as follows:

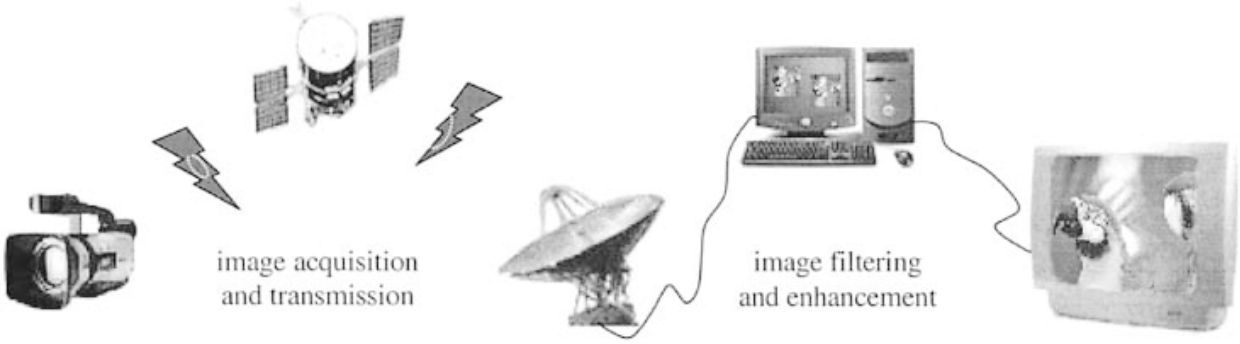
```
import cv2import numpyimg =
cv2.imread('pout.jpg')img_to_yuv =
cv2.cvtColor(img,cv2.COLOR_BGR2YUV)img_to_yuv[:,:,0]
=
cv2.equalizeHist(img_to_yuv[:,:,0])hist_equalization_result
= cv2.cvtColor(img_to_yuv,
cv2.COLOR_YUV2BGR) cv2.imwrite('result.jpg',hist_equalization_result)
```

14.7 COLOR IMAGE FILTERING

Image filtering changes the range (i.e. the pixel values) of an image, so the colors of the image are altered without changing the pixel positions, while image warping changes the domain (i.e. the pixel positions) of an image, where points are mapped to other points without changing the colors. *Color image filtering and enhancement refer to the process of noise reduction in the color image and enhancement of the visual quality of the image input.*

Noise reduces the perceptual quality of an image and thus limits the performance of the imaging system. The generation of high quality color images which are visually pleasing is of great importance in a variety of application areas. That pre supposes image filtering, since images captured with sensing devices and transmitted through communication networks are often corrupted by noise. Therefore, both filtering and enhancement constitute an important part of any image processing pipeline where the final image is utilized for visual inspection or for automatic analysis.

Figure 14.10



- **RGB Color Image Filters**

  Here we will focus on the most popular filters used to filter RGB color image. There are three different methods are used to filtering RGB color images.

  - **Average filter** Replace each pixel by the average of pixels in a square window surrounding this pixel.
  - **Median filter** Taking the median value instead of the average or weighted average of pixels in the window Median: sort all the pixels in an increasing order, take the middle one.
  - **Adaptive median filter** The adaptive median filtering has been introduced as an improvement to the standard median filtering, As we have explained before that the Median filtering can detect the noise but in the same it can't differentiate between the fine details and the noise. So, the main idea in the Adaptive Median Filter is to perform a spatial processing to determine which pixels in an image have been affected by impulse noise, and run the filter only in this pixel. The Adaptive Median Filter classifies pixels as noise by comparing each pixel in the image to its surrounding neighbor pixels. The size of the neighborhood is adjustable, as well as the threshold for the comparison. A pixel that is different from a majority of its neighbors, as well as being not structurally aligned with those pixels to which it is similar, is labeled as impulse noise. These noise pixels are then replaced by the median pixel value of the pixels in the neighborhood that have passed the noise labeling test.



CHECK YOUR PROGRESS

Q 1: What is the equation used to convert RGB pixel to HSI color format?

.....

.....

Q 2: What is PMF?

.....

.....

Q 3: What is CDF?

.....

.....

Q 4: What are the uses of histogram?

.....

.....

Q 5: What are the different types of filters?

.....

.....



14.8 LET US SUM UP

- Primarily color depends on the reflectance properties of an object. We see only those rays that are reflected, while others are absorbed.
- The human visual system can distinguish hundreds of thousands of different color shades and intensities, but only around 100 shades of grey.
- Colors are described with hue, saturation, intensity, brightness.
- Wavelength of visible colors occurs between about 400nm (violet) and 700nm (red) on the electromagnetic spectrum
- The human retina has 3 kinds of cones and the response of each type of cone as a function of the wavelength.
- The tristimulus theory of color perception implies that any color can

be formed from a mix of the three primary colors as red, green and blue, but some colors cannot be matched in this way. If one of the primaries is added to one of these unmatchable colors, it may be considered to have a negative weighting of that particular primary.

- Color models provide a standard way to specify a particular color, by defining a 3D coordinate system, and a subspace of color.
- There are different colour models such as RGB, H SI, CMY and YIQ.
- An image histogram, shows the frequency of pixels intensity values.
- In an image histogram, the x axis shows the gray level intensities and the y axis shows the frequency of these intensities.
- Probability Mass Function and Cumulative Distributive Function are both used in histogram equalization as the first and the second step.
- Histogram equalization enhances the contrast and brightness and give the unrealistic effects.
- Filtering changes the in range of the image without altering the position.
- There are three different types of filter, namely average, median, Adaptive median.



14.9 FURTHER READING

1. Gonzales, R. C., & Woods, R. E. (2008). Digital image processing. Prentice Hall.
2. Jayaraman, S. Esakkirajandan, S., & Veerakumar, T. (2009). Digital Image Processing, New Delhi: Tata McGraw-Hill Education Private Limited.

14.10 ANSWERS TO THE CHECK YOUR PROGRESS

Ans to Q No 1: Visible colors occur between about 400nm (violet) and 700nm (red) on the electromagnetic spectrum.

Ans to Q No 2: Hue, Saturation and Intensity.

Ans to Q No 3: The Tristimulus theory of color perception seems to imply that any color can be formed from a mix of the three primary colors as red, green and blue, but some colors can be matched in this way.

Ans to Q No 4: CIE defined three standard primaries, called *X*, *Y* and *Z* that can be added to form all visible colors. The primary *Y* was chosen so that its color matching function exactly matches the luminous-efficiency function for the human eye. X and Y gives the normalized amount of color. And Z is based on z=1-x-y.

Ans to Q No 5: RGB model consider that an image consists of three independent image planes, one in each of the primary colors: red, green and blue. Specifying a particular color is by specifying the amount of each of the primary components present.

Ans to Q No 6: The CMY (cyan-magenta-yellow) model is a *subtractive* model appropriate to absorption of colors. The CMY model asks what is subtracted from white.

Ans to Q No 7: I=1/2(R+G+B).

Ans to Q No 8: PMF stands for probability mass function. It gives the probability of each number in the data set or you can say that it basically gives the count or frequency of each element.

Ans to Q No 9: CDF stands for cumulative distributive function. It is a function that calculates the cumulative sum of all the values that are calculated by PMF. It basically sums the previous one.

Ans to Q No 10: Histogram use for analysis the image, for brightness and equalize the image.

Ans to Q No 11: Average filter, Median filter, Adaptive Median filter.



14.11 MODEL QUESTIONS

Q 1: Described the color fundamentals.

Q 2: Write about the human perception of color.

Q 3: Explain RGB color model.

Q 4: Explain CMY color model.

Q 5: Explain HSI color model.

Q 6: Explain YIQ color model.

Q 7: Discuss the procedure Gray Level to Color Transformation.

Q 8: What is histogram? Why a it necessary form color image?

Q 9: Write about the filtering of color image.

*** ***** ***

UNIT 15 : WAVELET BASED IMAGE PROCESSING

UNIT TRUCTURE

15.1 Learning objectives

15.2 Introduction

15.3 Wavelet

15.4 Wavelet Transform

15.4.1 Continuous Wavelet Transform

15.4.2 2D Continuous Wavelet Transform

15.5 Discrete Wavelet Transform

15.6 Haar Wavelet

15.7 Filter Bank

15.8 Embedded Image Coding

15.9 Let Us Sum Up

15.10 Further Reading

15.11 Answers to Check Your Progress

15.12 Model Questions

15.1 LEARNING OBJECTIVES

After going through this unit, you will be able to :

- learn about wavelet and wavelet transform
- describe the types of wavelet transform
- explain application of wavelet transforms
- understand filter bank, embedded image coding

15.2 INTRODUCTION

Wavelet Transform is an efficient tool to represent an image. The wavelet transform allows multi-resolution analysis of an image. The aim of the transform is to extract relevant information from an image. Wavelet transforms are the most powerful and widely used techniques in the field of image processing. A wavelet transform divides a signal into a number of segments, each corresponding to a different frequency band. Fourier Transform is also a powerful technique for image analysis. It gives

information about frequency content of a signal. The major disadvantage of Fourier transform is that it has only frequency resolution and no time resolution. This means that although we might be able to determine all the frequencies present in a signal, we do not know when they are present. To overcome this problem, in the past decades several multi-resolution solutions have been developed which are more or less able to represent a signal in the time and frequency domain at the same time.Multi-resolution methods attempt to find a specific frequency at a specific location, which is the main disadvantage of Fourier Transform and Short-Time Fourier Transform. However, it is not possible to find a specific frequency at a specific location simultaneously. Therefore, as a tradeoff between time-frequency representations, multi-resolution methods are created.

15.3 WAVELET

A wave is an oscillating function of time or space that is periodic. The wave is an infinite length continuous function in time or space. In contrast, wavelets are localized waves. A wavelet is a waveform of an effectively limited duration that has an average value of zero.In other words, wavelet is a mathematical function used for dividing a given function or continuous-time signal into different scale components. Usually one can assign a frequency range to each scale component.A function ψ (x) can be called a wavelet if it processes the following properties.

a. The function integrates to zero, or equivalently, its Fourier transform denoted as (w) is zero at the origin:

(15.1)

This implies $\psi(w)\big|_{w=0} = 0$ in the frequency domain

b. It is square integrable or equivalently, has finite energy:

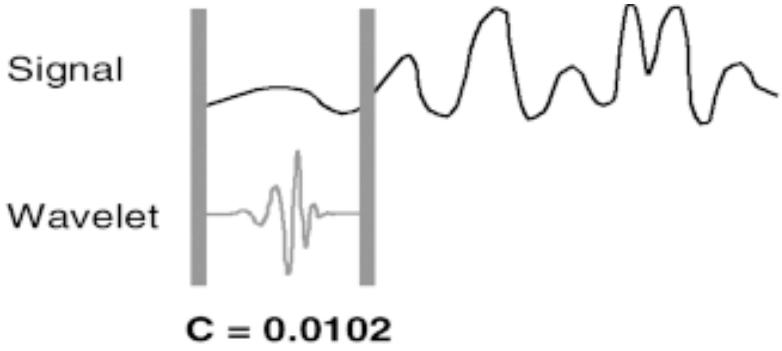
$$\int_{-\infty}^{\infty} |\psi(x)|^2 dx < \infty \qquad (15.2)$$

c. The Fourier transform (x) must satisfy the admissibility condition given by

(15.3)

Equation (15.1) suggests that the function is either oscillatory or has a wavy appearance. Equation (15.2) implies that most of the energy in is confined to finite interval. Ideally, the function is exactly zero outside the finite interval. This implies that the function is a compactly supported function. Equation (15.3) is useful in formulating the inverse wavelet transform. An example of signal and wavelet has be given in Figure 15.1

Figure 15.1: Example of Signal and wavelet



$$C\psi = \int_{-\infty}^{\infty} \frac{|\psi)w)|^2}{|w|} c$$

15.4 WAVELET TRANSFORM

The wavelet transform(WT) provides a time and frequency representation of a signal. The very small wavelet suggests that the function has to be well localized. Fourier basis functions are localized in frequency, but not in time. So, Fourier analysis is the ideal tool for the efficient representation of very smooth, stationary signals. Wavelet basis functions are localized in time and frequency. So, wavelet analysis is anideal tool for representing signal that contains discontinuities (in the signal or its derivatives) or for signals that are not stationary.The wavelet transform maps a function of one dimension into a two- dimensional picture. Using this idea, a mathematical microscope known as multi-resolution analysis will permit us to look at the details that are added as we go from one scale to another. This transform has excellent signal compaction properties for many classes of real- world signals while being

computationally very efficient. Therefore, it has been applied to almost all technical fields including image compression, denoising,numerical integration, and pattern recognition.

15.1.1 Continuous Wavelet Transform

In Continuous Wavelet Transform continuous shift and scale parameters are considered. The 1-D Continuous Wavelet Transform (CWT), also called the integral wavelet transform (IWT) of a signal, can be defined as follows

$$W_f(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\alpha}^{\alpha} x(t)\, \psi^{*}\left(\frac{t-b}{a}\right) dt$$

The continuous wavelet transform is a function of two variables and . Here is the scaling parameter and is the shifting parameter. is the mother wavelet or the basis function and it is used as a prototype for generating all the basis function. In mathematics, a basis function is an element of a particular basis for a function space. Every continuous function in the function space can be represented as a linear combination of basis functions, just as every vector in a vector space can be represented as a linear combination of basis vectors. The translation parameter or the shifting parameter b gives the time information in the wavelet transform. It indicates the location of the window and it is shifted through the signal. The scale parameter gives the frequency information in the wavelet transform. The wavelet transform was defined generally without the need to specify aparticular mother wavelet Ø. The user can select or design mother wavelet Ø according to application needs.The inverse 1D continuous wavelet transform is given by

$$x(t) = \frac{1}{C\psi} \int_{0}^{\alpha} \int_{-\alpha}^{\alpha} W_f(a,b)\, \psi_{a,b}\, db\, \frac{da}{a^2}$$

15.2.1 2D Continuous Wavelet Transform

The two-dimensional continuous wavelet transform (2D CWT) is a tool to analyse non-stationary signal. The2D CWT provides localized spectral information of the analyzed signal. It also has the advantage over the2D discrete wavelet transform (DWT) in that it covers the domain of the analyzed data with a continuousanalysis from which detailed, shift-invariant spectral information of different positions and orientations canbe obtained. The two-dimensional CWT (2D CWT) is an extension of the one-dimensional CWT, with the translation parameter being a vector in (x,y) plane. It can be generally defined as

$$W_f(a,b,\theta) = < f, \psi_{a,b,\theta} \text{ and } \Psi_{a,b,\theta(x)=a^{-1}\psi\left\{R_\theta^{-1}\left\{\frac{x-b}{a}\right\}\right\}}$$

where is a two dimensional square integrable real space function, is the wavelet mother function, is the scalar parameter, is the position vector, is the standard rotation matrix of rotating angle , and the angle brakets (< ….>) denotes the scalar product of the Hilbert space.

15.5 DISCRETE WAVELET TRANSFORM

The Discrete Wavelet Transform (DWT) is obtained by filtering the signal through a series of digital filters at different scales. If a signal is denoted by s(n) where n=0,1,2,……, M-1 then the forward discrete wavelet transform can be defined as

$$W_\phi(j_0,k) = \frac{1}{\sqrt{M}} \sum_n S(n)\, \phi_{j_0,k}(n) \qquad (15.1)$$

$$W_\Psi(j,k) = \frac{1}{\sqrt{M}} \sum_n S(n)\, \Psi_{j,k}(n) \qquad (15.2)$$

15.6 HAAR WAVELET

Haar wavelet is a sequence of rescaled “square-shaped” functions which together form a wavelet family or basis. The Haar sequence is

now recognised as the first known wavelet basis and extensively used as a teaching example. Haar sequence was proposed in 1909 by Alfréd Haar. The study of wavelets, and even the term “wavelet”, did not come until much later. As a special case of the Daubechies wavelet, the Haar wavelet is also known as Db1. The Haar wavelet is also the simplest possible wavelet. The technical disadvantage of the Haar wavelet is that it is not continuous, and therefore, not differentiable. This property can, however, be an advantage for the analysis of signals with sudden transitions, such as monitoring of tool failure in machines.

15.7 FILTER BANK

In signal processing, a filter bank is an array of band-passfilters that separates the input signal into multiple components, each one carrying a single frequencysub-band of the original signal. One application of a filter bank is a graphic equalizer, which can attenuate the components differently and recombine them into a modified version of the original signal. The process of decomposition performed by the filter bank is called analysis (meaning analysis of the signal in terms of its components in each sub-band); the output of analysis is referred to as a subband signal with as many subbands as there are filters in the filter bank. The reconstruction process is called synthesis.

In digital signal processing, the term filter bank is also commonly applied to a bank of receivers. The difference is that receivers also down-convert the subbands to a low center frequency that can be re-sampled at a reduced rate. The same result can sometimes be achieved by undersampling the bandpass subbands.

Another application of filter banks is signal compression when some frequencies are more important than others. After decomposition, the important frequencies can be coded with a fine resolution. Small differences at these frequencies are significant and a coding scheme that preserves these differences must be used. On the other hand, less important frequencies do not have to be exact.

The vocoder uses a filter bank to determine the amplitude information of the subbands of a modulator signal (such as a speech) and uses them to control the amplitude of the subbands of a carrier signal (such as the output of a guitar or synthesizer), thus imposing the dynamic characteristics of the modulator on the carrier.

15.8 EMBEDDED IMAGE CODING

The general principle of embedded image coding is that each successive bit of the bit stream that is received reduces the distortion of the reconstructed image by a certain amount. In order to achieve this, the information in the bit stream has to be organized in a decreasing order of importance. Embedded coding supports progressive image transmission by the property of generating the bits in the bit stream in their relative order of importance. A given image coded at a certain bit rate in an embedded fashion stores all the lower rode codes at the beginning of the bit stream. The coding process can be stopped before or when the target bit rate is met. Similarly, the decoder can interrupt decoding at any point in the bit stream and can reconstruct any lower rate image. Thus, progressive transmission is supported by transmitting a coarser version of the image first, followed, progressively, by the refinement details. The option of lossless reconstruction is still retained if the complete bit stream is generated, transmitted and decoded. The embedded image coders are built upon three major components: (i) wavelet transform, (ii) quantization scheme and (iii) significance-map encoding.



CHECK YOUR PROGRESS

Q 1: The transform which posseses the multi-resolution property is _____

Q 2: A function that integrate to 0 is called _____

Q 3: What is the relation between scale and frequency?

.....

.....



15.9 LET US SUM UP

- A wavelet is an oscillatory functions of finite duration.
- The continuous wavelet transform represents a signal as a function of time and scale.
- Discrete Wavelet Transform (DWT) can be implemented through (a) filter bank or (b) lifting scheme
- Multi-resolution analysis is possible with wavelet transform.
- Wavelets are widely used in the field of image compression, image denoising and image watermarking.



15.13 FURTHER READING

1. Gonzales, R. C., & Woods, R. E. (2008). Digital image processing. Prentice Hall.
2. Jayaraman, S. (2009). S. Esakkirajandan T. Veerakumar. Digital Image Processing, New Delhi: Tata McGraw-Hill Education Private Limited.

15.11 ANSWER TO CHECK YOUR PROGRESS

Ans to Q No 1: Wavelet

Ans to Q No 2: Wavelet

Ans to Q No 3: A higher scale corresponds to a more stretched wavelet and a low scale corresponds to a compressed wavelet. This implies that when the scale is high, the changes are slow and the frequency is low. When the scale is low , the changes are rapid and the frequency is high.



15.12 MODEL QUESTIONS

Q 1: What is wavelet and wavelet transform?

Q 2: What is the difference between CWT and DWT.

Q 3: Write the equation for DWT.

*** ***** ***



REFERENCES (For All Units of the Course)

1. Gonzales, R. C., & Woods, R. E. (2008). Digital image processing. Prentice Hall.
2. Jayaraman, S. (2009). S. Esakkirajandan T. Veerakumar. Digital Image Processing, New Delhi: Tata McGraw-Hill Education Private Limited.
3. Easton, R. L. (2010). Fundamentals of digital image processing.
4. Annadurai, S. (2007). Fundamentals of digital image processing. Pearson Education India.
5. Gunturk, B. K., & Li, X. (Eds.). (2012). Image restoration: fundamentals and advances. CRC Press.
6. Katsaggelos, A. K. (2012), Digital Image Restoration (Springer Series in Information Sciences), Springer
7. Salazar, N. (2015). Developments and Applications of Image Restoration. Clanrye International.
8. Joshi, M. A. (2018). Digital image processing: An algorithmic approach. PHI Learning Pvt. Ltd.

https://in.mathworks.com