Chapter 2

BACKGROUND

2.1. Image Processing

2.1.1. Digital Image Definitions

Vision

Vision allows humans to perceive and understand the world surrounding them. Computer vision aims to duplicate the effect of human vision by electronically perceiving and understanding an image. Giving computers the ability to see is not an easy task - we live in a three dimensional (3D) world, and when computers try to analyze objects in 3D space, available visual sensors (e.g., TV cameras) usually give two dimensional (2D) images, and this projection to a lower number of dimensions incurs an enormous loss of information. In order to simplify the task of computer vision understanding, two levels are usually distinguished; low-level image processing and high level image understanding.

Low-level methods usually use very little knowledge about the content of images. High level processing is based on knowledge, goals, and plans of how to achieve those goals. Artificial intelligence (AI) methods are used in many cases. High-level computer vision tries to imitate human cognition and the ability to make decisions according to the information contained in the image.
Digital Image

A digital image $a[m,n]$ described in a 2D discrete space is derived from an analog image $a(x,y)$ in a 2D continuous space through a sampling process that is frequently referred to as digitization. The 2D continuous image $a(x,y)$ is divided into $N$ rows and $M$ columns. The intersection of a row and a column is termed a pixel [AKJ1998]. The value assigned to the integer coordinates $[m,n]$ with $\{m=0,1,2,...,M-1\}$ and $\{n=0,1,2,...,N-1\}$ is $a[m,n]$. In fact, in most cases $a(x,y)$ which we might consider to be the physical signal that impinges on the face of a 2D sensor is actually a function of many variables including depth, color, and time.

The value assigned to every pixel is the average brightness in the pixel rounded to the nearest integer value. The process of representing the amplitude of the 2D signal at a given coordinate as an integer value with $L$ different gray levels is usually referred to as amplitude quantization or simply quantization.

Displays

The displays used for image processing, particularly the display systems used with computers have a number of characteristics that help determine the quality of the final image. They are refresh rate, interlacing and resolution.
**Refresh Rate**

The Refresh rate is defined as the number of complete images that are written to the screen per second. For standard video the refresh rate is fixed at the values of either 29.97 or 25 images. For computer displays the refresh rate can vary with the common values being 67 images and 75 images. At values above 60 images visual flicker is negligible at virtually all illumination levels.

**Interlacing**

To prevent the appearance of visual flicker at refresh rates below 60 images, the display can be interlaced. Standard interlace for video systems is 2:1. Since interlacing is not necessary at refresh rates above 60 images, an interlace of 1:1 is used with such systems. In other words, lines are drawn in an ordinary sequential fashion: \(1,2,3,4,\ldots,N\).

**Resolution**

The pixels stored in computer memory, although they are derived from regions of finite area in the original scene, may be thought of as mathematical points having no physical extent. When displayed, the space between the points must be filled in. This generally happens as a result of the finite spot size of a cathode-ray tube (CRT). The brightness profile of a CRT spot is approximately
Gaussian and the number of spots that can be resolved on the display depends on the quality of the system. It is relatively straightforward to obtain display systems with a resolution of 72 spots per inch (28.3 spots per cm.) This number corresponds to standard printing conventions. If printing is not a consideration, then higher resolutions, in excess of 30 spots per cm are attainable.

2.1.2. Low-level digital image processing

Low-level computer vision techniques overlap almost completely with digital image processing, which has been practiced for decades. The sequences of processing steps commonly recognized are given in figure 2.1.2.1 [RGRW1992].

![Image Processing Steps](image.png)
Image Acquisition

An image is captured by a sensor (such as a TV camera) and digitized. We can obtain digital images by conversion of analog images (such as 35mm prints, slides, transparencies or reflective art) into digital images with a scanner, or else by directly capturing the object or scene into digital form by means of a digital camera or video-capturing device. Some devices are listed below in Table 2.1.2.1

Table 2.1.2.1 Image Acquisition Devices

<table>
<thead>
<tr>
<th>Sl. No</th>
<th>Name of Device</th>
<th>Specialization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Digital Camera</td>
<td>Pictures can be instantly reviewed from the camera's LCD panel, and then re-taken if not satisfactory.</td>
</tr>
<tr>
<td>2</td>
<td>Flat Bed Scanners</td>
<td>Product images of approximately 640 pixels on the longest side and at least 16 bit color depth can be produced and saved in JPEG format.</td>
</tr>
<tr>
<td>3</td>
<td>Complementary metal oxide semiconductor (CMOS) image sensor</td>
<td>Digital images captured can be sorted, edited, corrected, and prepared for presentation, either electronically or through print media, with the accompanying software</td>
</tr>
<tr>
<td>Sl. No</td>
<td>Name of Device</td>
<td>Specialization</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>4</td>
<td>TV Camera or Vidicon Tube</td>
<td>Image can be focused onto a photoconductive target. Electric current produces as the beam passes over target. Current proportional to the intensity of light at each point.</td>
</tr>
<tr>
<td>5</td>
<td>Charged Coupled Device (CCD) Camera</td>
<td>This consists of an array of photosensitive cells. Each cell produces an electric current dependent on the incident light falling on it. There is less geometric distortion and more linear video output.</td>
</tr>
</tbody>
</table>

**Pre-processing**

Pre-processing is a common name for operations with images at the lowest level of abstraction, both input and output are intensity images. The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions or enhances some image features important for further processing.

Four categories of image pre-processing methods, according to the size of the pixel neighborhood are used for the calculation of new pixel brightness.
1. Pixel brightness transformations
2. Geometric transformations
3. Pre-processing methods that use a local neighborhood of the processed pixel, and
4. Image restoration that requires knowledge about the entire image.

Other classifications of image pre-processing methods exist. They are explained in Appendix A. Image pre-processing methods use the considerable redundancy in images. Redundancy implies repetitions of noise values which are not used for further processing. Neighboring pixels corresponding to one object in real images have essentially the same or similar brightness value. Thus, distorted pixel can often be restored as an average value of neighboring pixels. If pre-processing aims at correcting some degradation in the image, the nature of a priori information is important.

1. Knowledge about the nature of the degradation; only very general properties of the degradation are assumed.

2. Knowledge about the properties of the image acquisition device, and conditions under which the image was obtained. The nature of noise (usually its spectral characteristics) is sometimes known.

3. Knowledge about objects that are searched for in the image, which may simplify the pre-processing very considerably.
If knowledge about objects is not available in advance it can be estimated during the processing.

**Pixel brightness transformations**

Brightness transformations modify pixel brightness. The transformation depends on the properties of a pixel itself.

*Brightness correction*

This transformation considers original brightness pixel position in the image. The other transformations are only based on the original value.

*Gray scale transformation*

Gray scale transformation changes brightness without regard to position in the image and transformation is necessary based on the gray level value in the image.
Position dependent brightness correction

Ideally, the sensitivity of image acquisition and digitization devices should not depend on position in the image, but this assumption is not valid in many practical cases.

Geometric transformations

Geometric transformations permit the elimination of geometric distortion that occurs when an image is captured. An example is an attempt to match remotely sensed images of the same area taken after one year, when the more recent image was probably not taken from precisely the same position. To inspect changes over the year, it is necessary first to execute a geometric transformation, and then subtract one image from the other. A geometric transformation is a vector function \( T \) that maps the pixel \((x,y)\) to a new position \((x',y')\). The transformation equations are either known in advance or can be determined from known original and transformed images. Several pixels in both images with known correspondence are used to derive the unknown transformation.

A geometric transformation consists of two basic steps

1. Determining the pixel co-ordinate transformation.
Mapping of the co-ordinates of the input image pixel to the point in the output image is known as pixel coordinate transformation. The output point co-ordinates should be computed as continuous values (real numbers) as the position does not necessarily match the digital grid after the transform.

2. Finding the point in the digital raster, this matches the transformed point and determining its brightness.

Brightness is usually computed as an interpolation of the brightness's of several points in the neighborhood.

**Pixel co-ordinate transformations**

General cases of finding the co-ordinates of a point in the output image after a geometric transform is usually approximated by a polynomial equation.

This transform is linear with respect to the coefficients $a_k, b_k$, which are coordinates of the image. If pairs of corresponding points $(x,y), (x',y')$ in both images are known, it is possible to determine $a_k, b_k$ by solving a set of linear equations.
Nearest neighbor interpolation

Nearest neighbor interpolation Assigns to the point (x,y) the brightness value of the nearest point g in the discrete raster [RGRW1992]. There are various types of neighborhood connectivity of pixels are there they are

1. 4 Connectivity.
2. 8 Connectivity.
3. N Connectivity

Image segmentation

A computer tries to separate objects from the image background using image processing techniques. Object description and classification in a totally segmented image is also understood as a part of low-level image processing. A central problem, called segmentation is to distinguish objects from background. For intensity images (ie, those represented by point-wise intensity levels) the four popular approaches are: threshold techniques, edge-based methods, region-based techniques, and connectivity-preserving relaxation methods [RGRW1992].

Threshold techniques, which make decisions based on local pixel information, are effective when the intensity levels of the objects fall squarely outside the range of levels in the background. As spatial information is ignored, blurred region boundaries can create destruction.
Edge-based methods center around contour detection: their weakness in connecting together broken contour lines make them prone to failure in the presence of blurring.

A region-based method usually proceeds as follows: the image is partitioned into connected regions by grouping neighboring pixels of similar intensity levels. Adjacent regions are then merged under some criterion involving perhaps homogeneity or sharpness of region boundaries. Over stringent criteria create fragmentation; lenient ones overlook blurred boundaries and over merge. Hybrid techniques using a mix of the above methods are also popular.

A connectivity-preserving relaxation-based segmentation method, usually referred to as the active contour model, was proposed recently. The main idea is to start with some initial boundary shapes represented in the form of spline curves, and iteratively modify it by applying various shrink or expansion operations according to some energy function.

Image representation and description

Description, also called feature selection, deals with extracting features of interest. In character recognition description considered are holes and bays. It is to differentiate one part of the alphabet from another.
Image recognition and interpretation

The aim of shape recognition is to make the computer recognize particular shapes or patterns in a big source bitmap.

Interpretation and analysis of remote sensing imagery involves the identification and / or measurement of various targets in an image in order to extract useful information about them. Targets in remote sensing images may be any feature or object which can be observed in an image, and have the following characteristics:

- Targets may be a point, line, or area feature. This means that they can have any form from fish or zooplankton in ocean, from a bridge or roadway, to a large expanse of water or a field.
- The target must be distinguishable; it must contrast with other features around it in the image.

Knowledge base

Knowledge about a problem domain is coded into an image processing system in the form of a knowledge base. The various elements of image processing will use the knowledge base for doing the processing. The image acquired by means of image acquisition device is stored in the knowledge base. This image is preprocessed to remove noise. The noise-removed image is
segmented, described and interpreted from the information available in the knowledgebase.

2.1.3. Noise

Images acquired through modern sensors may be contaminated by a variety of noise sources. Noise we refers to stochastic variations as opposed to deterministic distortions such as shading or lack of focus. It is assumed for this section that images formed from light using modern electro-optics are dealt with. In particular it is assumed that modern charge-coupled device (CCD) cameras are used. Nevertheless, most of the observations made in this context about noise and its various sources hold equally well for other imaging modalities.

While modern technology has made it possible to reduce the noise levels associated with various electro-optical devices to almost negligible levels, one noise source can never be eliminated and thus forms the limiting case when all other noise sources are "eliminated". The various types of noise are listed below

1. Photon Noise
2. Thermal Noise
3. On-chip Electronic Noise
4. KTC Noise
5. Amplifier Noise
6. Quantization Noise

Photon Noise

When the observed physical signal is based upon light, then the quantum nature of light plays a significant role. A single photon at 500 nm carries an energy of \( E = h \nu = 3.97 \times 10^{-19} \) joules. Modern CCD cameras are sensitive enough to be able to count individual photons. The noise problem arises from the fundamentally statistical nature of photon production. We cannot assume that, in a given pixel for two consecutive but independent observation intervals of length \( T \), the same number of photons will be counted. Photon production is governed by the laws of quantum physics which restrict about an average number of photons within a given observation window.

Thermal Noise

An additional, stochastic source of electrons in a CCD well is thermal energy. Electrons can be freed from the CCD material itself through thermal vibration and then, trapped in the CCD well, and be indistinguishable from "true" photoelectrons. By cooling the CCD chip it is possible to reduce significantly the number of "thermal electrons" that give rise to thermal noise or dark current. As the integration time \( t \) increases, the number of thermal electrons increases. The probability distribution of thermal electrons is also a Poisson process where the
rate parameter is an increasing function of temperature. There are alternative techniques (to cooling) for suppressing dark current and these usually involve estimating the average dark current for the given integration time and then subtracting this value from the CCD pixel values before the A/D converter. While this does reduce the dark current average, it does not reduce the dark current standard deviation and it also reduces the possible dynamic range of the signal.

**On-chip Electronic Noise**

This noise originates in the process of reading the signal from the sensor, in this case, through the field effect transistor (FET) of a CCD chip. This noise has some impedance on the image.

**KTC Noise**

Noise associated with the gate capacitor of an FET is termed KTC noise and can be non-negligible. The output RMS value of this noise voltage is given by:

\[
\text{KTC noise (voltage)} = \sqrt{\frac{2kT}{C}}
\]

where \( C \) is the FET gate switch capacitance, \( k \) is Boltzmann's constant, and \( T \) is the absolute temperature of the CCD chip measured in K. Using the relationships, the output RMS value of the KTC noise is expressed in terms of the number of photoelectrons.
Amplifier Noise

The standard model for this type of noise is additive, Gaussian, and independent of the signal. In modern well-designed electronics, amplifier noise is generally negligible. The most common exception to this is in color cameras where more amplification is used in the blue color channel than in the green channel or red channel leading to more noise in the blue channel.

Quantization Noise

Quantization noise is inherent in the amplitude quantization process and occurs in the analog-to-digital converter, ADC. The noise is additive and independent of the signal when the number of levels $L \geq 16$. This is equivalent to $B \geq 4$ bits. For a signal that has been converted to electrical form and thus has a minimum and maximum electrical value, an appropriate formula is used for determining the SNR.
2.2. Remote Sensing

2.2.1. Concept of Remote Sensing

Remote Sensing is defined as the science and technology by which the characteristics of objects of interest can be identified, measured or analyzed without direct contact.

Electro-magnetic radiation that is reflected or emitted from an object is the usual source of remote sensing data. However any media such as gravity or magnetic fields can be utilized in remote sensing.

A device to detect the electro-magnetic radiation reflected or emitted from an object is called a "remote sensor" or "sensor". Cameras and scanners are examples of remote sensors.

A vehicle to carry the sensor is called a "platform". Aircraft or satellites are used as platforms.

The technical term "remote sensing" was first used in the United States in the 1960's, and encompassed photogrammetry, photo-interpretation, photo-geology etc. Since Landsat-1, the first earth observation satellite that was launched in 1972 remote sensing has become widely used.
The characteristics of an object can be determined, using reflected or emitted electro-magnetic radiation from the object. Each object has a unique and different characteristic of reflection or emission depending upon the type of object or the environmental condition. “Remote sensing is a science and technology to identify and understand the object or the environmental condition through the uniqueness of the reflection or emission”.

The remote sensing data will be processed automatically by computer and / or manually interpreted by humans, and finally utilized in agriculture, land use, forestry, geology, hydrology, oceanography, meteorology, environment etc [SLH2000].

**Characteristics of Electro-Magnetic Radiation**

**Electro-magnetic radiation** is a carrier of electro-magnetic energy by transmitting the oscillation of the electro-magnetic field through space or matter. The transmission of electro-magnetic radiation is derived from Maxwell equations. Electro-magnetic radiation has the characteristics of both wave motion and particle motion.

1. Characteristics as wave motion : Electro-magnetic radiation can be considered as a transverse wave with an electric field and a magnetic field. A plane wave has its electric field and magnetic field in the perpendicular plane to the transmission direction. The two fields are located at right
angles to each other. The **wavelength** \( \lambda \), **frequency** \( \nu \) and velocity \( v \) have the following relation.

\[
\lambda = \frac{v}{\nu}
\]

Electro-magnetic radiation is transmitted in a vacuum of free space with the velocity of light \( c = 2.998 \times 10^8 \) m/sec and in the atmosphere with a reduced velocity but similar to that in a vacuum. The frequency \( \nu \) is expressed as a unit of hertz (Hz), which is the number of waves which are transmitted in a second.

2. Characteristics as particle motion: Electro-magnetic radiation can be treated as a photon or a light quantum. The energy \( E \) is expressed as follows.

\[
E = h\nu
\]

where \( h \): Plank's constant \( \nu \): frequency

The photoelectric effect can be explained by considering electro-magnetic radiation as composed of particles. Electro-magnetic radiation has four elements: **frequency** (or wavelength), **transmission direction**, **amplitude** and **plane of polarization**. The amplitude is the magnitude of oscillating electric field. The square of the amplitude is proportional to the energy transmitted by electro-magnetic radiation. The energy radiated from an object is called radiant energy. A
plane including the electric field is called the plane of polarization. When the plane of polarization forms a uniform plane, it is called linear polarization.

The four elements of electro-magnetic radiation are related to different information contents. Frequency (or wavelength) corresponds to the color of an object in the visible region, which is given by a unique characteristic curve relating to the wavelength and the radiant energy. In the microwave region, information about objects is obtained using the Doppler shift effect in frequency that is generated by a relative motion between an object and a platform. The spatial location and shape of objects are given by the linearity of the transmission direction, as well as by the amplitude. The plane of polarization is influenced by the geometric shape of objects in the case of reflection or scattering in the microwave region. In the case of radar, horizontal polarization and vertical polarization have different responses on a radar image.

Types of Remote Sensing with Respect to Wavelength Regions

Remote sensing is classified into three types with respect to the wavelength regions: (1) Visible and Reflective Infrared Remote Sensing, (2) Thermal Infrared Remote Sensing and (3) Microwave Remote Sensing

The energy source used in the visible and reflective infrared remote sensing is the sun. The sun radiates electro-magnetic energy with a peak
wavelength of 0.5 μm. Remote sensing data obtained in the visible and reflective infrared regions mainly depends on the **reflectance** of objects on the ground surface. Therefore, information about objects can be obtained from spectral reflectance. However, laser radar is exceptional because it does not use solar energy but uses the laser energy of the sensor.

The source of radiant energy used in thermal infrared remote sensing is the object itself, because any object with a normal temperature (the temperature when the body is in a stable state) will emit electro-magnetic radiation with a peak at about 10 μm.

In the microwave region, there are two types of microwave remote sensing: passive microwave remote sensing and active microwave remote sensing. In passive microwave remote sensing, the **microwave radiation** emitted from an object is detected, while **back scattering coefficient** is detected in active microwave remote sensing.

**Definition of Radiometry**

In remote sensing, electro-magnetic energy reflected or emitted from objects is measured. The measurement is based on either **radiometry** or **photometry**, with different technical terms and physical units.
Radiometry is used for physical measurement of a wide range of radiation from x-ray to radio wave, while photometry corresponds to the human perception of visible light based on the human eye's sensitivity.

**Radiant energy** is defined as the energy carried by electro-magnetic radiation and expressed in the unit of joule (J). **Radiant flux** is radiant energy transmitted as a radial direction per unit time and expressed in the unit of watt (W). **Radiant intensity** is radiant flux radiated from a point source per unit solid angle in a radiant direction and expressed in the unit of Wsr⁻¹. **Irradiance** is radiant flux incident upon a surface per unit area and expressed in the unit of Wm⁻². **Radiant emittance** is radiant flux radiated from a surface per unit area, and expressed in a unit of Wm⁻². **Radiance** is radiant intensity per unit-projected area in a radial direction and expressed in the unit of Wm⁻² sr⁻¹.

**Reflectance**

**Reflectance** is defined as the ratio of incident flux on a sample surface to reflected flux from the surface. Reflectance ranges from 0 to 1. Reflectance was originally defined as a ratio of incident flux of white light to reflected flux in a hemisphere direction. The equipment used to measure reflectance is called a spectrometer.
**Albedo** is defined as the reflectance using the incident light source from the sun. **Reflectance factor** is sometimes used as the ratio of reflected flux from a sample surface to reflected flux from a perfectly diffuse surface. Reflectance with respect to wavelength is called **spectral reflectance**. A basic assumption in remote sensing is that spectral reflectance is unique and different from one object to an unlike object.

Reflectance with a specified incident and reflected direction of electromagnetic radiation or light is called **directional reflectance**. The two directions of incident and reflection can be directional, conical or hemispherical by making nine possible combinations.

### 2.2.2. Sensors

**Classification of Sensors**

It is expected that some new types of sensors will be developed in the future. **Passive sensors** detect the reflected or emitted electro-magnetic radiation from natural sources, while **active sensors** detect reflected responses from objects which are irradiated from artificially generated energy sources, such as radar. Each is divided further into **non-scanning** and **scanning systems**.
A sensor classified as a combination of passive, non-scanning and non-imaging methods is a type of profile recorder, for example a microwave radiometer. A sensor classified as passive, non-scanning and imaging methods, is a camera, such as an aerial survey camera or a space camera, for example the one on board the Russian COSMOS satellite.

Sensors classified as a combination of passive, scanning and imaging are classified further into image plane scanning sensors, such as TV cameras and solid state scanners, and object plane scanning sensors, such as multispectral scanners (optical-mechanical scanner) and scanning microwave radiometers.

Those sensors, which use lenses in the visible and reflective infrared region, are called optical sensors.

Characteristics of Optical Sensors Radiation

Optical sensors are characterized by spectral, radiometric and geometric performance. The spectral characteristics are spectral band and bandwidth, central wavelength, response sensitivity at the edges of band, spectral sensitivity at outer wavelengths and sensitivity of polarization.

The sensitivity of film and the transmittance of the filter, and nature of the lens characterize the sensors using film. Scanner type sensors are specified by the
spectral characteristics of the detector and the spectral splitter. In addition, chromatic aberration is an influential factor. The **radiometric characteristics** of optical sensors are specified by the change of electro-magnetic radiation which passes through an optical system. They are radiometry of the sensor, sensitivity to **noise equivalent power**, **dynamic range**, signal to noise ratio (S/N ratio) and other noises, including quantification noise.

The geometric characteristics are specified by those geometric factors such as field of view (FOV), instantaneous field of views (IFOV), band-to-band registration, geometric distortion and alignment of optical elements.

IFOV is defined as the angle contained by the minimum area that can be detected by a scanner type sensor. For example in the case of an IFOV of 2.5 milli radians, the detected area on the ground will be 2.5 meters x 2.5 meters, if the altitude of sensor is 1,000 m above ground.

**Characteristics of Optical Detectors**

An element that converts the electro-magnetic energy to an electric signal is called a detector. There are various types of detectors with respect to the detecting wavelength.
Detectors are of three types: photoemission type, optical excitation type, and thermal effect type. Phototube and photo multiplier tubes are examples of the photoemission type, which has sensitivity in the region from ultra violet to visible light.

Photodiode, phototransistor, photoconductive detectors and linear array sensors are examples of optical excitation types, which have sensitivity in the infrared region. Photodiode detectors utilize electric voltage from the excitation of electrons, while phototransistor and photoconductive detector utilize electric current.

Thermocouple barometers and pyroelectric barometers are examples of the thermal effect type, which has sensitivity from near infrared to far infrared regions. However the response is not very high because of the thermal effect.

Detectivity denoted as D* (termed as D star) is usually related to sensitivity, expressed as NEP (noise equivalent power). D* is used for comparison between different detectors. NEP is defined as the signal input identical to the noise output. NEP depends on the type of detector, surface of detector or band of frequency. D* is inversely proportional to NEP.
2.2.3. Remote Sensing Devices

Cameras for Remote Sensing

Aerial survey cameras, multispectral cameras, panoramic cameras etc. are used for remote sensing.

Aerial survey cameras, sometimes called metric cameras are usually used on board aircraft or spacecraft for topographic mapping by taking stereo photographs with overlap. A typical aerial survey camera is RMK made by Carl Zeiss or RC series made by Leica Company.

Typical well known, examples of space cameras, are the Metric Camera on board the Space Shuttle by ESA, the Large Format Camera also on board the Space Shuttle by NASA, and the KFA 1000 on board COSMOS sent by Russia.

As the metric camera is designed for very accurate measurement of topography, the following requirements in optics as well as geometry should be specified and fulfilled.

1. Lens distortion should be minimal.
2. Lens resolution should be high and the image should be very sharp even in the corners.
3. Geometric relation between the frame and the optical axis should be established.

4. Lens axes and film planes should be vertical to each other.

5. A vacuum pressure plate should maintain film flatness.

6. Focal length should be measured and calibrated accurately.

7. Successive photographs should be made with high-speed shutter and film winding system.

8. To prevent the image motion of high speed moving objects during shutter time Forward Motion Compensation (FMC) should be used, particularly in the case of space cameras.

Multispectral cameras with several separate film scenes in the visible and reflective IR are mainly used for photo-interpretation of land surface covers.

Panoramic cameras are used for reconnaissance surveys, surveillance of electric transmission lines, supplementary photography with thermal imagery etc., because the field of view is so wide.

Film for Remote Sensing

Various types of films are used in cameras for remote sensing. Film can record the electromagnetic energy reflected from objects in the form of optical density in an emulsion placed on a film base of polyester. There are
panchromatic (black and white film), infrared film, color film, color infrared film, etc.

The spectral sensitivity of film is different depending on the film type. Black and white infrared film has wider sensitivity up to near infrared as compared with panchromatic film. Color film has three different spectral sensitivities according to three layers of primary color emulsion (B.G. R). Color infrared film has sensitivity up to 900 nm. Kodak aerial color film SO-242 has high resolution and is specially ordered for high altitude photography.

Generally film is composed of a photographic emulsion, which records various gray levels from white to black according to the reflectivity of objects.

A curve that shows the relationship between the exposure E and the photographic density is called the "characteristic curve". Usually the horizontal axis of the curve is log E, while the vertical axis is D (density) which is given as follows.

\[ D = \log \left( \frac{1}{T} \right) \text{ where } T : \text{transparency of film} \]
**Imaging Spectrometer**

**Imaging spectrometers** are characterized by a multispectral scanner with a very large number of channels (64-256 channels) with very narrow bandwidths, though the basic scheme is almost the same as an optical mechanical scanner or push broom scanner.

The optical systems of imaging spectrometers are classified into three types: diopptic system, dio and catopptic system which are adopted depending on the scanning system. In the case of object plane scanning, the catopptic system is best because the linearity of the optical axis is very good due to the narrow view angle and the observation wave range is so wide. However in the case of image plane scanning, the dioptic system or dioptic and catopptic system is best suited because the view angle should be wider.

As imaging spectrometer provides multiband imagery with a narrow wavelength range, is useful for rock type classification and ocean color analysis.

**Laser Radar**

A device which measures the physical characteristics such as distance, density, velocity, shape etc., using scattering, returned time, intensity, frequency and / or polarization of light called optical sensor. However as the actual light
used by the optical sensor is mostly laser, it is usually called **laser radar** or **lidar** (light detection and ranging).

Laser radar is an active sensor which is used to measure air pollution, physical characteristics of atmospheric constituents in the stratosphere and its spatial distribution. The theory of laser radar is also utilized to measure distance and it is called laser distance meter or laser altimeter for this application.

The main measurement object is the atmosphere although laser radar is also used to measure water depth, thickness of oil film or vividness of chlorophyll in vegetation.

In laser radar received light is converted to an electric signal that is displayed or recorded after A/D conversion. The effective distance of lidar depends on the relationship between the received light intensity and the noise level.

Lidar can be classified with respect to its physical characteristics, interactive effects, physical quantities etc.

Fluorescence lidar, Roman lidar and differential absorption lidar are utilized for measurement of density of gaseous bodies, while Doppler lidar is used
for measurement of velocity. The polarization effect of lidar is utilized for measurement of shapes.

There are several display modes for example, a scope with horizontal axis of distance and with vertical axis of intensity. PPI (plane position indication) with gray level in polar coordinate system, RHI (range high indication) with a display of the vertical profile and THI (time height indication) with a horizontal axis of elapsed time and with a vertical axis of altitude.

*Characteristics of Radar Image*

The main objective of microwave remote sensing is to estimate the property of objects by interpreting the features of the radar image. Typical objects to be measured by microwave remote sensing are mountainous landforms, subsurface geology, sea wind and waves etc. In order to estimate these properties, it is very important to understand the effects of microwave backscattering on the objects.

Two factors of microwave characteristics are of importance: frequency (or wavelength) and polarization. In microwave remote sensing, various wavelengths (or frequency) such as L band, C band, X band, P band etc. will be used ranging from millimeter wavelengths (1 mm - 1cm) up to about 30 cm. According to the wavelength or frequency, secular reflection will occur, so that
the surface roughness can be detected if multi-frequency radar images are compared.

Polarization is defined as the oscillating direction involved in an electric field. Usually transmitted microwave and received microwave will have a choice between horizontal polarization and vertical polarization. Therefore four combinations: HH, HV, VH and VV can be used for SAR. The backscattering characteristics are also different with respect to polarization.

In future, SAR systems with functions of multi-frequency and multi-polarization will be onboard earth observation satellites.

**Image Reconstruction of SAR**

Raw data from SAR are records of backscattering in time sequence that are returned from the ground targets. Return signal from a point P is recorded in the expanded range in the range direction, which is identical to the pulse width. In addition, the return signal from a point P is also expanded in the azimuth direction because point P continues to be radiated by microwave pulses during the flight motion.
Data processing to generate an image of gray tone corresponding to the backscattering intensity of each point on the ground is called image reconstruction of synthetic aperture radar (SAR).

Image reconstruction is divided into range compression and azimuth compression that makes compression of expanded signals in both range and azimuth directions into a point signal. Adopting Fourier transformation to achieve convolution of received signals and a reference function usually carries out the compression.

The reference function of range compression is the complex conjugate of the transmitted signal, while the reference function of azimuth compression is a complex conjugate of the modulated signal by chirp modulation.

The slant range to a point on the ground is expressed as the quadratic function of time with respect to the movement of the platform. The change of the slant range is called range migration. The first order term is called range walk resulting from the earth rotation, while the second order term is called range curvature.

In image reconstruction, there is a major problem, called speckle, which is due to high frequency noise. In order to reduce the speckle, mulch-look processing is applied in which range compression and azimuth compression with
respect to subdivided frequency domains are independently overlaid three or four times, termed the number of looks. Sometimes a median filter or local averaging may be applied to reduce the speckle. The speckle will be reduced by the square root of the number of looks, although the spatial resolution declines in proportion to the number of looks.

**Microwave Radiometer**

Microwave radiometers or passive type microwave sensors are used to measure the thermal radiation of the ground surface and/or atmospheric condition.

Reyleigh-Jean's law expresses brightness temperature measured by a microwave radiometer, which is the resultant energy of thermal radiation from the ground surface and the atmospheric media. Multi-channel radiometers with multi-polarization are used to avoid the influences of unnecessary factors to measure the specific physical parameter.

The simplest radiometer is the total power radiometer. This system has a mixer to enable it to mix high frequency of a local oscillator in order to amplify the high signal after transforming to a low frequency. However the influence of system gain variation cannot be neglected in this system.
The Dicke radiometer can reduce the influence of system gain variation by introducing a switch generator which allows it to receive the antenna signal and noise source of constant temperature, alternatively of which antenna signal can be detected later on, synchronously with the switch generator.

The zero-balance Dicke radiometer can reduce the influence of system gain variation and increase the sensitivity further by adding a noise generator to the Dicke radiometer in order to increase the sensitivity about two times higher than total power radiometer.

Some of the remote sensing devices are listed table 2.2.3.1

Table 2.2.3.1. Remote Sensing Devices

<table>
<thead>
<tr>
<th>Sl. No</th>
<th>Name of Device</th>
<th>Specialization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Radar</td>
<td>Detect the amount of that energy which is radiated by the object</td>
</tr>
<tr>
<td>2</td>
<td>Lidar (Laser Imaging raDAR)</td>
<td>Operate in the ultraviolet, visible and near infrared wavelengths</td>
</tr>
<tr>
<td>3</td>
<td>Microwave Radiometer</td>
<td>Detect reflected responses from objects that are irradiated from artificially-generated energy sources</td>
</tr>
<tr>
<td>Sl. No</td>
<td>Name of Device</td>
<td>Specialization</td>
</tr>
<tr>
<td>-------</td>
<td>---------------------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>4</td>
<td>Infra-Red Line Scanner</td>
<td>Can show, by the heat signature on the ground (IRLS)</td>
</tr>
<tr>
<td>5</td>
<td>Sonar systems</td>
<td>Artificially produced energy to a target and record the reflected component</td>
</tr>
<tr>
<td>6</td>
<td>Frame Camera</td>
<td>It creates an image by using a lens to receive the electromagnetic energy from the ground and transmit it to the film where it is recorded</td>
</tr>
</tbody>
</table>
2.3 Image Classification

2.3.1. Classification Steps

Image Classification of Remote Sensing

Identification of features by remote sensing is effective for global assessment of geometric characteristics and general appraisal of ground cover types. Quantitative analysis using the computer enables identification of pixels based upon their numerical properties. It is also possible to count pixels for area estimates by quantitative analysis.

Image classification is the process of creating thematic maps from satellite imagery. A thematic map is an informational representation of an image, which shows the spatial distribution of a particular theme (class). Themes can be as diversified as their areas of interest. Examples of themes are soil, vegetation, water depth, fish, rock, zooplankton, phytoplankton and atmosphere. Inside a theme, there can be defined sub themes, and thus the process of classification needs to be very refined.

Feature Extraction

This is an optional step on the classification process, which serves only as a low level pre-processing of the image to reduce its spectral, or spatial dimensionality. It can be accomplished by using any type of spatial filters or
spectral transforms to reduce the data and/or enhance its multispectral features, or even by simply selecting a subset of bands. In this stage, the multispectral image is transformed into a feature image. An example of a feature extraction is the use of a Principal Components Stretch (PCS) to enhance the spectral characteristics of the image, which helps the analyst identify class similarities.

**Training of the Pixels**

In this step, pixels from the image are extracted to train the classifier to recognize patterns, which help differentiate the classes. Based on these patterns, the classifier creates discriminate functions to assign each pixel to a class in the feature space. The training of the pixels can be either Supervised or Unsupervised.

**Supervised Training:** It assumes some pre-determined knowledge about the spatial distribution of the classes on the image. The training points for each site are selected prior to the application of discriminate functions in the K-D feature space. Themes are known at the time when the classes are labeled. Classes are labeled before clustering.

**Unsupervised Training:** Classes are separated without any previous assumptions about their spatial distribution on the image.
Labeling

In this final stage of the image classification process, the discriminate functions are used to label all the pixels in the entire feature image. If the training of the pixels was supervised, then a previous knowledge of the class’s spatial distribution would allow the labeling of classes to be carried out upon the application of the discriminate functions to the feature space. If the training was unsupervised, the assignment of a label to each discriminated class would be subject to the analyst’s own labeling criteria.

2.3.2. Classification Procedures

Color pixels classification algorithm

Before classifying the player pixels of a current color image, the player pixels are extracted. Then, we transform the R, G and B features of each player pixel \( P(x,y) \) into the Hybrid Color Space features. In order to classify a player pixel \( P(x,y) \) of a color image, we consider the set of player pixels falling into a neighborhood of \( P(x,y) \).

The size of this neighborhood depends on the mean player’s size in the images. Then, for each player pixel, we evaluate a mean vector \( M_p = [m_{p1}, m_{p2}, m_{p3}]^T \) of the HCS features of the player pixels which belong to the neighborhood.
For each class $C_j$, we evaluate the Euclidian distance $D_j(x,y)$ between the mean vector $M_j$ of the class $C_j$ and the mean vector $M_P$ in the HCS space:

$$D_j(x,y) = \| M_j - M_P \| = \left( \sum_{k=1}^{K} (m_j^k - m_P^k)^2 \right)^{1/2}$$

Finally, a minimum decision rule is used to assign $P(x,y)$ to the class $C_j$ ($j=1,...,N_c$) for which $D_j(x,y)$ is minimum [RGRW1992].

A player pixel is classified with its own HCS features and the neighbor’s HCS features. This classification scheme achieves a selective smoothing of the image, with a window size that is the mean size of a player. Indeed, the neighborhoods, evaluated for each player pixel, don’t always contain the same number of player pixels. So, this smoothing procedure is achieved with a variable number of player pixels. It permits as to label a majority of the player pixels, which represent the same player to the same class, even if these pixels don’t form a uniform color region in the image. Indeed, a player wears a soccer suit, which can be composed of several colors.
PNN Image Classification

PNN stands for "probabilistic neural network". The PNN image classification algorithm compares each pixel of data with several known pixels from each of several classes. The algorithm then has a sense of how "close" a pixel is to each class and chooses the class for the pixel based on which class it was closest to.

K-means clustering

There are several variants of the k-means clustering algorithm, but most variants involve an iterative scheme that operates over a fixed number of clusters, while attempting to satisfy the following properties:

1. Each class has a center, which is the mean position of all the samples in that class.
2. Each sample is in the class whose center it is closest to it.

The basic k-means algorithm consists of the following steps:
Initialize

Loop until termination condition is met:

1. Assign each pixel in the image to a class such that the distance from this pixel to the center of that class is minimized.
2. For each class, recalculate the means of the class based on the pixels that belong to that class.

End loop

Distance Measure

The mean of the k-means algorithm is calculating the distance between each pixel and each class center. There are different distance measures that can be used. The most common are:

- L1 distance (Manhattan distance) is the absolute value of the component-wise difference between the pixel and the class. This is the simplest distance to calculate and may be more robust to outliers.
- L2 distance (Euclidean distance) is the square root of the component-wise square of the difference between the pixel and the class. Since only the results are compared the square root can be omitted. Computing the L2 distance
requires squaring the data, which introduces extra bits of precision into the data. The squaring operation is expensive in hardware. One advantage of this metric is that the distance is a sphere around the centroid.

Unsupervised Classification

Unsupervised Classification is performed most often by using clustering methods to assign each pixel in an image to spectral classes, of which the user has no foreknowledge. These procedures can be used to determine the number and location of the spectral classes into which the pixels are assigned. Using the existing information from ground maps, aerial photos, or ground visits, the output classes are identified.

Supervised Classification

Supervised classification procedures are the essential tools used to extract quantitative information from remotely sensed data. To perform it, one first determines the classes obtained from the satellite images. Some examples of class are high-density housing, medium density housing, water, fish, rock, zooplankton, phytoplankton and open space. For each class, a sample of pixels that correspond
to it is selected to allow a reasonable estimate of the range of pixels in each class. These ranges, called training sites, are saved in a vector file, which is then used to create a signature or spectral response pattern for each class. These signatures are then used to classify the full image by determining the most likely class for each individual pixel in the image.

2.3.3. Classification Techniques

Classification of remotely sensed data is used to assign corresponding levels with respect to groups with homogeneous characteristics, with the aim of discriminating multiple objects from each other within the image. The level is called class. Classification will be executed on the base of spectral or spectrally defined features such as density, texture etc. in the feature space. It can be said that classification divides the feature space into several classes based on a decision rule. In many cases, classification will be undertaken using a computer, with the use of mathematical classification techniques. Classification will be made according to the following procedures.

Step 1: Definition of Classification Classes. Depending on the objective and the characteristics of the image data, the classification classes should be clearly defined.
Step 2: **Selection of Features.** Features to discriminate between the classes should be established using multi-spectral and/or multi-temporal characteristics, textures etc.

Step 3: **Sampling of Training Data.** Training data should be sampled in order to determine appropriate decision rules. Classification techniques such as supervised or unsupervised learning will then be selected on the basis of the training data sets.

Step 4: **Estimation of Universal Statistics.** Various classification techniques will be compared with the training data, so that an appropriate decision rule is selected for subsequent classification.

Step 5: **Classification.** Depending upon the decision rule, all pixels are classified as a single class. There are two methods of pixel-by-pixel classification and per-field classification, with respect to segmented areas.

Popular techniques are listed in the following table 2.3.3.1.
Table 2.3.3.1. Classification Techniques.

<table>
<thead>
<tr>
<th>Serial No.</th>
<th>Techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Multi-level slice classifier</td>
</tr>
<tr>
<td>2</td>
<td>Minimum distance classifier</td>
</tr>
<tr>
<td>3</td>
<td>Maximum likelihood classifier</td>
</tr>
<tr>
<td>4</td>
<td>Fuzzy set theory</td>
</tr>
<tr>
<td>5</td>
<td>Expert systems</td>
</tr>
</tbody>
</table>

Step 6: Verification of Results. The classified results should be checked and verified for their accuracy and reliability.

Estimation of Population Statistics

a. Supervised classification

In order to determine a decision rule for classification, it is necessary to know the spectral characteristics or features with respect to the population of each class. The spectral features can be measured using ground-based spectrometers. However due to atmospheric effects, direct use of spectral features measured on the ground are not always available. For this reason, sampling of training data from clearly identified training areas, corresponding to defined classes is usually
made for estimating the population statistics. This is called supervised classification. Statistically unbiased sampling of training data should be made in order to represent the population correctly.

a. **Unsupervised Classification.**

In the case where there is less information in an area to be classified, only the image characteristics are used as follows.

(1) Multiple groups, from randomly sampled data, will be mechanically divided into homogeneous spectral classes using a clustering technique.

(2) The clustered classes are then used for estimating the population statistics. This classification technique is called unsupervised classification.

Maximum likelihood estimation is the most popular method by which the population statistics such as mean and variance are estimated to maximize the probability or likelihood from a defined probability density function within the feature space [RGRW 1992].

In most cases, the probability density function is selected to be a multiple normal distribution. The multiple normal distribution gives the following maximum likelihood estimator.

\[
\text{Mean; } \mu_{ci} = \frac{1}{n} \sum_{j=1}^{n} X_{ij} \ (i = 1,2,...,m)
\]

Variance - covariance matrix
\[ \Sigma_e = \frac{1}{m} \sum_{i=1}^{m} (x_i - \mu_e) (x_i - \mu_e) \]

where \( m \): number of bands and \( n \): number of pixels

Before adopting the maximum likelihood classification, it should be checked to determine if the distribution of training data would fit the normal distribution or not.

**Clustering**

**Clustering** is a grouping of data with similar characteristics [RGRW1992]. Clustering is divided into hierarchical clustering and non-hierarchical clustering as follows.

a. Hierarchical Clustering.

The similarity of a cluster is evaluated using a "distance" measure. The minimum distance between clusters will give a merged cluster after repeated procedures from a starting point of pixel-wise clusters to a final limited number of clusters.

The distances to evaluate the similarity are selected from the following methods.
(1) **Nearest neighbor method**: Nearest neighbor with minimum distance will form a new merged cluster.

(2) **Furthest neighbor method**: Furthest neighbor with maximum distance will form a new merged cluster.

(3) **Centroid method**: Distance between the gravity centers of two clusters is evaluated for merging a new merged cluster.

(4) **Group average method**: Root mean square distance between all pairs of data within two different clusters is used for clustering.

(5) **Ward method**: Root mean square distance between the gravity center and each member is minimized.

b. Non-hierarchical Clustering.

At the initial stage, an arbitrary number of clusters should be temporarily chosen. The members belonging to each cluster will be checked by selected parameters or distance and relocated into the more appropriate clusters with higher separability. The ISODATA method and K-mean method are examples of non-hierarchical clustering.

The ISODATA method is composed of the following procedures.

(1) All members are relocated into the closest clusters by computing the distance between the member and the clusters.
(2) The center of gravity of all clusters is recalculated and the above procedure is repeated until convergence.

(3) If the number of clusters is within a certain specified number, and the distances between the clusters meet a prescribed threshold, the clustering is considered complete.

Parallelepiped Classifier

The parallelepiped classifier (often termed multi-level slicing) divides each axis of multi-spectral feature space. The decision region for each class is defined on the basis of the lowest and the highest value on each axis. The accuracy of classification depends on the selection of the lowest and highest values in consideration of the population statistics of each class. In this respect, it is most important that the distribution of population of each class is well understood.
The parallelepiped classifier is very simple and easy to understand schematically. In addition the computing time will be minimum, when compared to other classifiers.

However the accuracy will be low especially when the distribution in feature space has covariance or dependency with oblique axes. Orthogonalization should be undertaken using principal component analysis, for example, before adopting the parallelepiped classifier.
The **minimum distance classifier** is used to classify unknown image data as classes that minimize the distance between the unknown image data and the class in multi-feature space [RGRW1992]. The distance is defined as an index of similarity so that the minimum distance is identical to the maximum similarity.

The following distances are often used in this procedure.

![Figure 2.3.3.2 Concept of Minimum Distance Classifier](image)

(1) **Euclidian distance.**

\[
d_k^2 = (X - \mu_k)^\top (X - \mu_k)
\]

This is used in cases where the variances of the population classes are different from each other. The Euclidian distance is theoretically identical to the similarity index.
(2) Normalized Euclidian distance.

The Normalized Euclidian distance is proportional to the similarity index, in the case of difference variance.

\[ d_k^2 = (X - \mu_k)^T \sigma_k^{-1} (X - \mu_k) \]

(3) Mahalanobis distance.

In cases where there is correlation between the axes in feature space, the Mahalanobis distance with variance-covariance matrix is used.

\[ d_k^2 = (X - \mu_k)^T \Sigma_k^{-1} (X - \mu_k) \]

where \( X \) : vector of image data (n bands)

\[ X = [ x_1, x_2, ..., x_n ] \]

\( \mu_k \) : mean of the kth class \( \mu_k = [ m_1, m_2, ..., m_n ] \)

\( \sigma_k \) : variance matrix

\[ \sigma_k = \begin{pmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \sigma_{nn} \end{pmatrix} \]

\( \Sigma_k \) : variance-covariance matrix

\[ \Sigma_k = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix} \]
Maximum Likelihood Classifier

The **maximum likelihood classifier** is one of the most popular methods of classification in remote sensing, in which a pixel with the maximum likelihood is classified into the corresponding class. The likelihood $L_k$ is defined as the posterior probability of a pixel belonging to class $k$ [RGRW1992].

$$L_k = P(k/X) = \frac{P(k) * P(X/k)}{\sum P(i) * P(X/i)}.$$  

where $P(k)$: prior probability of class $k$. 

![Figure 2.3.3.3 Concept of Maximum Likelihood Method](image)

- Probability density
- Assumptions of multivariate normal distribution
- Classification in band 1
- Classification in band 2
- Likelihood belonging to class B: big
- Likelihood belonging to class A: small
P(X/k) : conditional probability to observe X from class k, or probability density function

Usually P(k) are assumed to be equal to each other and ΣP(i)*P(X/i) is also common to all classes. Therefore Lk depends on P(X/k) or the probability density function.

For mathematical reasons, a multivariate normal distribution is applied as the probability density function. In the case of normal distributions, the likelihood can be expressed as follows.

\[
L_k(X) = \frac{1}{(2\pi)^{n/2}|\Sigma_k|^{1/2}} \exp \left[ -\frac{1}{2} (X - \mu_k)^T \Sigma_k^{-1} (X - \mu_k) \right]
\]

where \( n \): number of bands.

X: image data of n bands.

\( L_k(X) \): likelihood of X belonging to class k.

\( \mu_k \): mean vector of class k.

\( \Sigma_k \): variance-covariance matrix of class k.

\( |\Sigma_k| \): determinant of \( \Sigma_k \)

In the case where the variance-covariance matrix is symmetric, the likelihood is the same as the Euclidian distance, while in case where the
determinants are equal to each other, the likelihood becomes the same as the
Mahalanobis distances.

The maximum likelihood method has an advantage from the viewpoint of
probability theory, but care must be taken with respect to the following items.

(1) Sufficient ground truth data should be sampled to allow estimation of the
mean vector and the variance-covariance matrix of population.

2) The inverse matrix of the variance-covariance matrix becomes unstable
in the case where there exists very high correlation between two bands or the
ground truth data are very homogeneous. In such cases, the number of bands
should be reduced by a principal component analysis.

(3) When the distribution of the population does not follow the normal
distribution, the maximum likelihood method cannot be applied.

Classification using an Expert System

Experts interpret remote sensing images with knowledge based on
experience. However computer assisted classification utilizes only very limited
expert knowledge. The **expert system**, therefore, is a problem solving system which supports expert knowledge in a computer based system.

The following two types of knowledge are required for an expert system in remote sensing.

(1) Knowledge about image analysis.

Procedures for image analysis can be made only with adequate knowledge about image processing and analysis. A feedback system should be introduced for checking and evaluating the objectives and the results.

(2) Knowledge about the objects to be analyzed.

Knowledge about the objects to be recognized or classified should be introduced in addition to the ordinary classification method. The fact that forest does not exist over 3,000 meters above sea level is one example of the type of knowledge that can be introduced.

**Decision Tree Classifier**

The **decision tree classifier** is a hierarchically based classifier, which compares the data with a range of properly selected features. The selection of
features is determined from an assessment of the spectral distributions of the classes. There is no generally established procedure. Therefore an expert should design each decision tree or set of rules. When a decision tree provides only two outcomes at each stage, the classifier is called a binary decision tree classifier (BDT).

Generally a group of classes will be classified into two groups with the highest separable with respect to a feature.

Features often used are as follows.

1) Spectral values.

2) An index that is computed from spectral values. For example, the vegetation index is a popular index.

3) Any arithmetic value such as addition, subtraction or rationing.

4) Principal components.

The advantages of the decision tree classifier are that computing time is less than that for the maximum likelihood classifier and by comparison the statistical errors are avoided. However the disadvantage is that accuracy depends fully on the design of the decision tree and the selected features.
2.4. Classification of Oceanographic Images

2.4.1. Oceanography

2.4.1.1. Definition

Oceanography (also called marine science) is the study of the earth's oceans and their interlinked ecosystems and chemical and physical processes. There are five major divisions within the science:

- Marine geology, including plate tectonics and other study of the ocean floor;
- Physical oceanography, which is concerned with the physical attributes of the ocean (such as its temperature, salinity structure and currents);
- Chemical oceanography (the study of the chemistry of the ocean);
- Biological oceanography (also sometimes considered a subset of marine biology), which is the study of the flora and fauna of the ocean;
- Meteorological oceanography, which is concerned with how the atmosphere and the ocean interact.
Oceanography is the study of the sea, embracing and integrating all knowledge pertaining to the sea and its physical boundaries, the chemistry and physics of seawater, and marine biology.

2.4.1.2. Oceanic resources

An enormous numbers of oceanic resources are available. Some of the resources and their characteristics are described below. The major categorization of resources are

1. Marine Mammals
2. Vertebrates
3. Echinoderm
4. Coelenterates
5. Mollusks
6. Crustaceans

1. Marine Mammals

The characteristics of marine mammals are

- Vertebrates (backbone)
2. Vertebrates

The characteristics of vertebrates are:

- Have a backbone
- Single pump hearts
- Lay eggs
- Have paired fins (two of a kind)
- Have scales
- Have a backbone
- Can have bony skeletons (most fresh and salt water fish)
- Can have cartilaginous skeletons (sharks and rays)

3. Echinoderm

The characteristics of echinoderms are
• Invertebrate (no backbone)

• Radically symmetrical (all parts arranged in a circle around the body / no right or left side

• Cells separate into organs (are more advanced than worms)

4. Coelenterates

The characteristics of coelenterates are:

• Invertebrates (no backbone)

• Radically symmetrical (all parts arranged in a circle around the body / no right or left side

• Having stinging cells called endoblasts containing poison

• Having a specialization that when the "trigger" (poison-filled, barbed thread) is touched the poison is discharged

• Trap and paralyze their prey

5. Mollusks

The characteristics of mollusks are:
• Invertebrates (no backbone)
• Soft unsegmented bodies enclosed in a shell
• Can have a bivalve (two shells, e.g. clams)
• Can have a univalve (one shell, e.g. snail)
• Can have a tubular siphon under the head with arms that have suckers (e.g. octopus)

6. Crustaceans

The characteristics of crustaceans are:

• Have exoskeletons (a skeleton outside of their bodies)
• Have two pairs of antennae
• Are closely related to the crayfish (anthropoids, but they have only one pair of antennae)
• Use gills

Phytoplankton (or algae) constitutes the base of marine food web [SLH2000].
2.4.1.3. Oceanic classification

Wetlands

In general terms, wetlands are lands where saturation with water is the dominant factor determining the nature of soil development and the types of plant and animal communities living in the soil and on its surface. The single feature that most wetlands share is soil or substrate that is at least periodically saturated with or covered by water. The water creates severe physiological problems for all plants and animals except those that are adapted for life in water or in saturated soil.

Deepwater Habitats

Deepwater habitats are permanently flooded lands lying below the deepwater boundary of wetlands. Deepwater habitats include environments where surface water is permanent and often deep, so that water, rather than air, is the principal medium within which the dominant organisms live, whether or not they are attached to the substrate. As in wetlands, the dominant plants are hydrophytes; however, the substrates are considered non-soil because the water is too deep to support emergent vegetation.
Wetlands and deepwater habitats are defined separately because traditionally the term wetland has not included deep permanent water; however, both must be considered in an ecological approach to classification. Five major Systems: Marine, Estuarine, Riverine, Lacustrine, and Palustrine are defined. The first four of these include both wetland and deepwater habitats; the Palustrine includes only wetland habitats [LVFE1998].

2.4.1.4. Indian Ocean

The Indian Ocean is the third largest of the world's five oceans (after the Pacific Ocean and the Atlantic Ocean, but larger than the Southern Ocean and Arctic Ocean). Four critically important access waterways are the Suez Canal (Egypt), Bab el Mandeb (Djibouti-Yemen), Strait of Hormuz (Iran-Oman), and Strait of Malacca (Indonesia-Malaysia).

The geographical detail specifies that the body of water between Africa, the Southern Ocean, Asia, and Australia. Total area is of 68.556 million sq km. It includes Andaman Sea, Arabian Sea, Bay of Bengal, Flores Sea, Great Australian Bight, Gulf of Aden, Gulf of Oman, Java Sea, Mozambique Channel, Persian Gulf, Red Sea, Savu Sea, Strait of Malacca, Timor Sea, and other tributary water bodies.
The surface is dominated by counterclockwise gyre (broad, circular system of currents) in the southern Indian Ocean and unique reversal of surface currents in the northern Indian Ocean. Low atmospheric pressure over southwest Asia from hot, rising, summer air results in the southwest monsoon and southwest-to-northeast winds and currents, while high pressure over northern Asia from cold, falling, winter air results in the northeast monsoon and northeast-to-southwest winds and currents; ocean floor is dominated by the Mid-Indian Ocean Ridge and subdivided by the Southeast Indian Ocean Ridge, Southwest Indian Ocean Ridge, and Ninetyeast Ridge.

Natural resources like oil and gas fields, fish, shrimp, sand and gravel aggregates, placer deposits, and poly metallic nodules. Occasional icebergs pose navigational hazard in southern regions. Endangered marine species include the dugong, seals, turtles, and whales; oil pollution in the Arabian Sea, Persian Gulf, and Red Sea.

In connection with economy the Indian Ocean provides major sea routes connecting the Middle East, Africa, and East Asia with Europe and the Americas. It carries a particularly heavy traffic of petroleum and petroleum products from the oilfields of the Persian Gulf and Indonesia. Its fish are of great and growing importance to the bordering countries for domestic consumption and export. Fishing fleets from Russia, Japan, South Korea, and Taiwan also exploit the Indian Ocean, mainly for shrimp and tuna. Large reserves of hydrocarbons are
being tapped in the offshore areas of Saudi Arabia, Iran, India, and western Australia. An estimated 40% of the world's offshore oil production comes from the Indian Ocean. Beach sands rich in heavy minerals and offshore placer deposits are actively exploited by bordering countries, particularly India, South Africa, Indonesia, Sri Lanka, and Thailand.

2.4.2. Location-wise Classification

2.4.2.1. Introduction

This deals with the classification of images location-wise. Digital images are represented as an array of values. These images are acquired by a remote sensing mechanism. The values represent the pixel information about the image. The pixel information's are the red, blue and green compositions.

Location-wise classification implies that when the coordinate values of the images are given the resources available in the surrounding area will be classified. An ocean occupies a vast area, and large resources are available in different locations. A certain area may be of importance for some important valuable resources. Oceanographic classification becomes very useful in recording the resources in a particular area.
The images acquired from the remote sensing devices are stored as x and y coordinate values in the format of BMP, GIF or JPEG format. The bitmap data includes the header information regarding the image size, height, width and color representations, and the pixel information about the image.

The pixel information is the basic color combination. This combination may be 8 bits or 16 bits. The bit representation depends on the resolution of the image. If the bit representation is high it means the resolution of the image is high. The coordinate values are represented as an array of information. From the array of information, the location-wise classification will be able to extract the needed information depending on the value of the coordinate specified.

2.4.2.2. Background

Location identification

The image is stored in the storage location as a one-dimensional array of bitmap pixel data. The array contains the pixel information as red, green and blue combinations. For a single pixel three locations are needed. The array of data representation is given below showing the index values for each pixel. Pixel-array is the array name for storing the image. Pixel-array[i] indicates the i\textsuperscript{th} location information.
The first pixel information represented in the array starts with index 0.

Pixel-array[0] - red band information for the first pixel.

The second pixel information represented in the array starts with index 3.


The third pixel information represented in the array starts with index 6.


Similarly all other pixel information is stored in the one-dimensional array. So the $n^{th}$ pixel information is available at location $3*n$. Comparing the pixel value with the already existing pixel values performs the classification process.
Representation of a 3 x 5 image in terms of x and y co-ordinate value

<table>
<thead>
<tr>
<th></th>
<th>0,0</th>
<th>0,1</th>
<th>0,2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4,0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Representation of a 3 x 5 image in terms of pixel number

<table>
<thead>
<tr>
<th>Pixel 1</th>
<th>Pixel 2</th>
<th>Pixel 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel 4</td>
<td>Pixel 5</td>
<td>Pixel 6</td>
</tr>
<tr>
<td>Pixel 7</td>
<td>Pixel 8</td>
<td>Pixel 9</td>
</tr>
<tr>
<td>Pixel 10</td>
<td>Pixel 11</td>
<td>Pixel 12</td>
</tr>
<tr>
<td>Pixel 13</td>
<td>Pixel 14</td>
<td>Pixel 15</td>
</tr>
</tbody>
</table>

The above two different representations of image can specify the organization of pixels in a 3 x 5 image. It can be known that pixel 1 is at location 0,0. Pixel 2 is at location 0,1 and so on. From the above matrix it can be seen that pixel 8 is at location 2,1.
Each pixel is stored as three bands of data in the bit map array. The three different bands of values are red, green and blue. So each pixel occupies three different locations. In some image representations they are stored as some number bands.

The purpose is to extract the pixels that are used for classifying the specific location of an image. The location of the pixel used for classification is identified from the array of image data. It is calculated by the equation

\[
\text{Pixel identification} = 3(j + n \times i)
\]

\[\text{twodimage}[i,j] = \text{onedimage}[\text{pixel identification}]\]

where \(\text{twodimage}\) represents the two dimensional representation of the image and \(\text{onedimage}\) represents the one dimensional array representation in the BMP format.

\(i\) represents the row coordinate value in the original image.

\(j\) represents the column coordinate value in the original image.

\(n\) represents the maximum number of columns in the original image.

For example supposing one wishes to classify the image of size 5 x 3. it specifies that one is going to classify the image that has 5 rows and 3 columns. It is assumed that the region interest for classification is 2,1. It indicates that the
pixel used for classification has the row (i) value 2 and column (j) value 1. It means the image is to be classified near the eighth pixel. The maximum number of columns (n) in this image is 3.

The array location for classification is identified as follows:

The value for j is 1.
The value for i is 2.
The value for n is 3.

Pixel identification = 3(j+ n * i)  
Pixel identification = 3(1+3*2)  
Pixel identification = 3(1 +6)  
Pixel identification = 3*7  
Pixel identification = 21  

From this result it is known that the pixel values used for classifying the image at co-ordinate value 2,1 starts at location 21 in the bit map array representation.

After the exact location is identified, the surrounding pixels are classified by means of classification algorithms. So the pixels have to be processed near the array location 21 for classifying the image.
Techniques

Classification of remotely sensed data is used to assign corresponding levels with respect to groups with homogeneous characteristics, with the aim of discriminating multiple objects from each other within the image. The techniques used for classification are Minimum distance classifier, Parallelepiped classifier and Maximum likelihood classifier.

2.4.2.3. Description

The row value and column value which indicate the location where classification is necessary are used to find the beginning and end of row value and beginning and end of column value for the purpose of classification. If the beginning of row is less than zero (top boundary of image) it is initialized as zero. Similarly if the beginning of column value is less than zero (left boundary of image) it is initialized as zero. If the end of row is greater than or equal to y resolution it is initialized as one less than y resolution. Similarly, if the end of column is greater than or equal to x resolution it is initialized as one less than x resolution.

The classification is performed only for the co-ordinates of row and column beginning and row and column ending. The classification is based on the
image classification techniques used. From the original image the pixels that match the particular class for selection are highlighted.

In order to classify the image we must know the pixel location in the bitmap pixel array. The pixel index (pixel identification) value for classification is calculated based on the row and column value selected for classification. It is calculated by the formula

\[ \text{Pixel identification} = 3(j + n \times i) \]

where \( i \) indicates the row selected for classification,

\( j \) indicates the column selected for classification and

\( n \) indicates the number of columns in the original image.

The pixels having the index calculated by the pixel identification variable are compared with the existing classes. These comparisons are performed based on the classification techniques available like parallelepiped classification technique, minimum distance technique, maximum classification technique, matching by difference technique or other classification techniques. Since only a small number of pixels are involved in classification rather than all the pixels in the image, the speed of the computation is increased. The scientist who is interested in finding the resources at a particular location uses this type of location-wise classification.
2.4.2.4. Flow Diagram

START

The row value and column value to know the location involved for classification are accepted.

Calculate the number of pixels in the image.

Set the beginning of row and column as well as end of row and column for indicating the location for classification using the row and column value specified.

A
The pixel index used for classification is calculated as \( \text{pidx} = (\text{column} + \text{xres} \times \text{row}) \).

For the selected location of row and column value assign the pixel array value to different bands.

Classify the image near the pixel index using a different color value (255).

Display the classified image to know the location involved in classification.

STOP
2.4.2.5. High Level Algorithm

Step 1: The row value and column value to know the location involved for classification are accepted.

Step 2: Calculate the number of pixels in the image.

Step 3: Set the beginning of row and column as well as end of row and column for indicating the location for classification using the row and column value specified.

Step 4: The pixel index used for classification is calculated as $\text{pidx} = 3 \times (\text{column}\times \text{res}\times \text{row})$.

Step 5: For the selected location of row and column value assign the pixel array value to different bands.

Step 6: Classify the image near the pixel index using a different color value (255).

Step 7: Display the classified image to know the location involved in classification.
2.4.2.6. Algorithm for location-wise classification

Algorithm `loccl(pixels[], rowval, colval)`

// pixels – the one dimensional pixel array that contains the pixel value of an image.
// rowval – the row location for classification.
// colval – the column location for classification.
// npix – number of pixels.

1. \( npix = (xres*(yres+1)) \);
2. \( \text{rowbegin} = \text{rowval}-50; \)
3. \( \text{rowend} = \text{rowval}; \)
4. \( \text{colbegin} = \text{colval}-50; \)
5. \( \text{colend} = \text{colval}; \)
6. if (rowbegin <= 0) rowbegin = 0;
7. if (rowend >= yres) rowend = yres-1;
8. if (colbegin <= 0) colbegin = 0;
9. if (colend >= xres) colend = xres-1;
10. \( k = 0; \)
11. for(\(i = \text{rowbegin; i<=rowend;i++}\))
12. for(j = colbegin; j<=colend;j++)

13. pidx = (j+ xres*i);
14. band[k][1] = pixels[pidx];
15. band[k][2] = pixels[pidx + 1];
16. band[k][3] = pixels[pidx + 2];
17. k = k + 1;
}
18. for( p = 0 ;p<npix-1 ;p++)
{
19. modipix[p] = 255;
}
20. Display the image.

The results of location-wise classification are used to find the resources by specifying the coordinates. The location calculating formulae are used to find the related information of pixels from the array of pixel values. The classification algorithms are used to find the classification of the resources. Depending on the type of classification algorithm used, the resources are clearly classified. The various types of algorithms are minimum distance classification algorithm, parallelepiped classification algorithm, maximum likelihood classification algorithm and matching by difference classification algorithm. The matching by difference classification is described in the subsequent chapter. This classification
is useful for the scientist to find the place where a large amount of valuable resources are available.

2.4.2.7. Complexity Analysis

The assignment of data in step 1 takes place in constant time.

The calculation of the number of pixel in the image in step 2 is also performed in a constant time.

The beginning of row and column as well as the end of row and column is set for indicating the location for classification in step 3. This takes place in a constant time.

The calculation of pixel index used for classification in step 4 depends on the row (r) value and column (c) value involved in the classification. So the computing time is O(rc).

Assigning the pixel array value to different bands in step 5 depends on row (r) value and column (c) value involved in the classification. So the computing time is O(rc).
Classifying the selected area of image in step 6 depends on the number of pixels involved \(3 \times r \times c\) in classification. So the computing time in this step is three times \(O(rc)\).

Displaying the classified image in step 7 depends on the number of pixels involved \(3 \times r \times c\) in classification. So the computing time in this step is three times \(O(rc)\).

### 2.4.2.7. Merits and Demerits

The location-wise classification is used to classify the image for a selected range of region. So a large number of computations are reduced for this classification process. In order to find the occurrence of resource in a particular location we need not classify the whole image.

The location-wise classification is combined with other classification techniques such as parallelepiped classification, minimum distance classification and matching by difference classification. This combination is used to classify the needed resource in a particular location.
2.4.2.8. Sample Result

Figure 2.4.2.1 shows the result of implementing classification using location-wise classification. In the following image class 1 indicates black area, class 2 indicates green area and class 3 indicates red area.

Figure 2.4.2.1. Classification using Location-wise Classification

Note: 100,100 denotes x-axis 100 pixels and y-axis 100 pixels.

White space represents the location used for classification.

The results of location-wise classification are used to find the resources by specifying the coordinates. The location calculating formulae are used to find the related information of pixels from the array of pixel values. The classification algorithms are used to find the classification of the resources in a precise way. This classification is helpful for the scientist to find the place where a large amount of valuable resources are available.
Table 2.4.2.1 shows the comparison of classification by location-wise classification in image img-org at location 100,100. Image classification using parallelepiped classification, minimum distance classification and matching by difference classification are specified in the table.

Total number of pixels under the selected location is 2601.

**Table 2.4.2.1. Pixel classified using Location-wise Classification**

<table>
<thead>
<tr>
<th>Serial No</th>
<th>Classification Technique</th>
<th>Pixels classified under class 1</th>
<th>Pixels classified under class 2</th>
<th>Pixels classified under class 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Parallelepiped Classifier</td>
<td>398</td>
<td>67</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>Minimum Distance Classifier</td>
<td>1407</td>
<td>813</td>
<td>381</td>
</tr>
<tr>
<td>3</td>
<td>Matching by Difference</td>
<td>Class 1 selected</td>
<td>Nil</td>
<td>Nil</td>
</tr>
</tbody>
</table>
From the above comparison it may be concluded that in the image named as img-org at location 100, 100 class 1 resource is abundantly available. The second large number of resources is class 2 and only a small number of class 3 resources are present.