GUDLAVALLERU ENGINEERING COLLEGE

(An Autonomous Institute with Permanent Affiliation to JNTUK, Kakinada)

Seshadri Rao Knowledge Village, Gudlavalleru – 521 356.

OPEN ELECTIVE

Learning Material

GEOINFORMATICS

UNIT –I

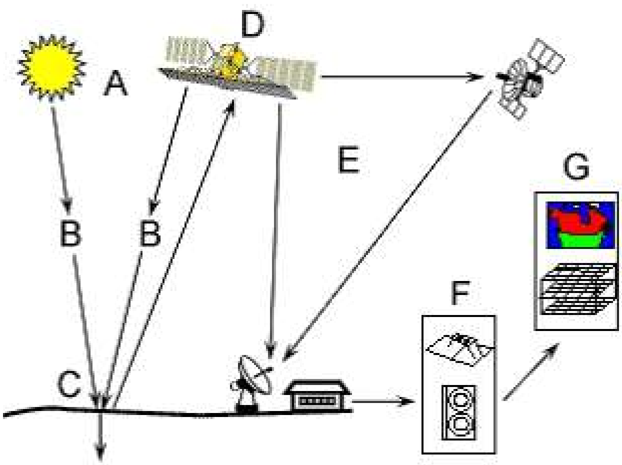
EMR and its Interaction with Atmosphere & Earth material

Definition of remote sensing and its components , Electromagnetic spectrum wavelength regions important to remote sensing, Wave theory, Particle theory, Stefan-Boltzmann and Wien's Displacement Law, Atmospheric scattering, Absorption , Atmospheric windows, spectral signature concepts, typical spectral reflective characteristics of water, vegetation and soil.

Learning Material

"**Remote sensing** is the science (and to some extent, art) of acquiring information about the Earth's surface without actually being in contact with it. This is done by sensing and recording reflected or emitted energy and processing, analysing, and applying that information."

In much of remote sensing, the process involves an interaction between incident radiation and the targets of interest. This is exemplified by the use of imaging systems where the following seven elements are involved. Note, however that remote sensing also involves the sensing of emitted energy and the use of non-imaging sensors.



COMPONENTS

1. Energy Source or Illumination (A) - the first requirement for remote sensing is to have an energy source which illuminates or provides electromagnetic energy to the target of interest.

2. Radiation and the Atmosphere (B) - as the energy travels from its source to the target, it will come in contact with and interact with the atmosphere it passes through. This interaction may take place a second time as the energy travels from the target to the sensor.

3. Interaction with the Target (C) - once the energy makes its way to the target through the atmosphere, it interacts with the target depending on the properties of both the target and the radiation.

4. Recording of Energy by the Sensor (D) - after the energy has been scattered by, or emitted from the target, we require a sensor (remote - not in contact with the target) to collect and record the electromagnetic radiation.

5. Transmission, Reception, and Processing (E) - the energy recorded by the sensor has to be transmitted, often in electronic form, to a receiving and processing station where the data are processed into an image (hardcopy and/or digital).

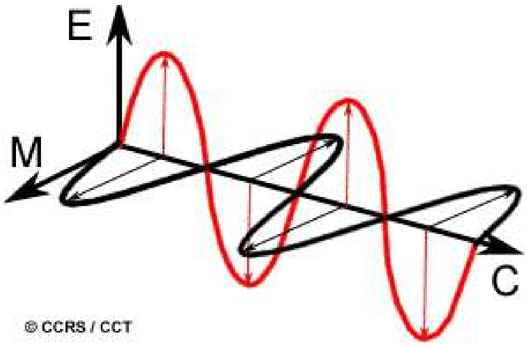
6. Interpretation and Analysis (F) - the processed image is interpreted, visually and/or digitally or electronically, to extract information about the target which was illuminated.

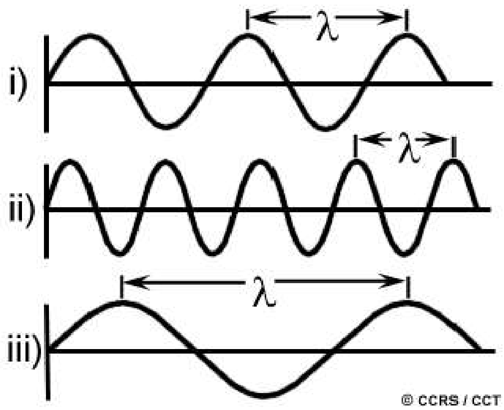
7. Application (G) - the final element of the remote sensing process is achieved when we apply the information we have been able to extract from the imagery about the target in order to better understand it, reveal some new information, or assist in solving a particular problem.

These seven elements comprise the remote sensing process from beginning to end.

Wave theory

Electromagnetic radiationconsists of an electrical field(E) which varies in magnitude in a direction perpendicular to the direction in which the radiation is travelling, and a magnetic field (M) oriented at right angles to the electrical field. Both these fields travel at the speed of light (C). Two characteristics of electromagnetic radiation are particularly important for understanding remote sensing





The relation between the wavelength (**λ**) and frequency (v) of EMR is based on the following formula

v=c/λwhere c= velocity of light.

Particle theory offers insight into how electromagnetic energy interacts with matter. It suggests that EMR is composed of many discrete units called photons/quanta.

The energy of photon is

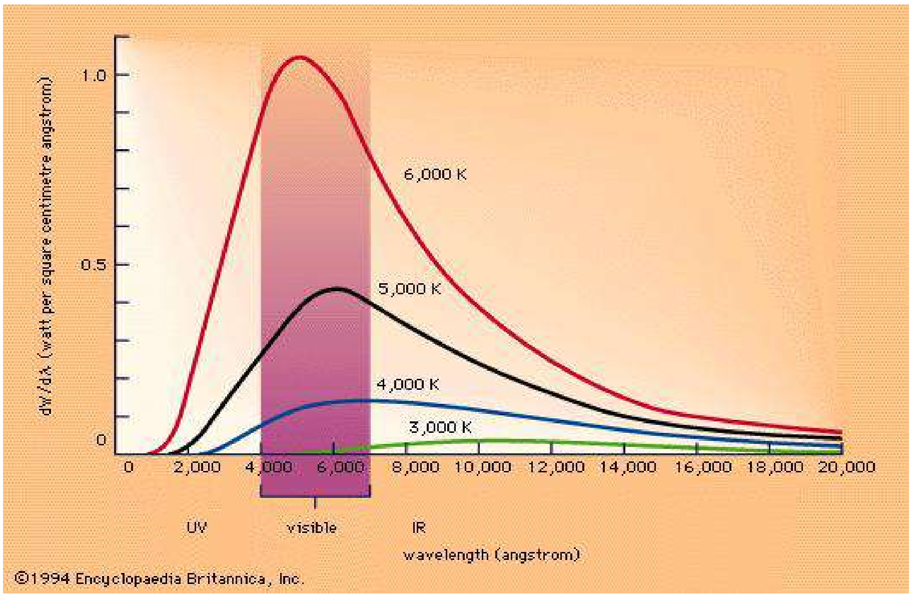
$$Q = hc / \lambda = h \nu$$

Where Q is the energy of quantum, h = Planck's constant

Stefan–Boltzmann law, statement that the total radiant heat energy emitted from a surface is proportional to the fourth power of its absolute temperature. Formulated in 1879 by Austrian physicistJosef Stefan as a result of his experimental studies, the same law was derived in 1884 by Austrian physicist Ludwig Boltzmann from thermodynamic considerations: if *E* is the radiant heat energy emitted from a unit area in one second and *T* is the absolute temperature (in degrees Kelvin), then $E = \sigma T^4$, the Greek letter sigma (σ) representing the constant of proportionality, called the Stefan–Boltzmann constant. This constant has the value 5.6704×10^{-8} watt per metre^2·K^4. The law applies only toblackbodies, theoretical surfaces that absorb all incident heat radiation.

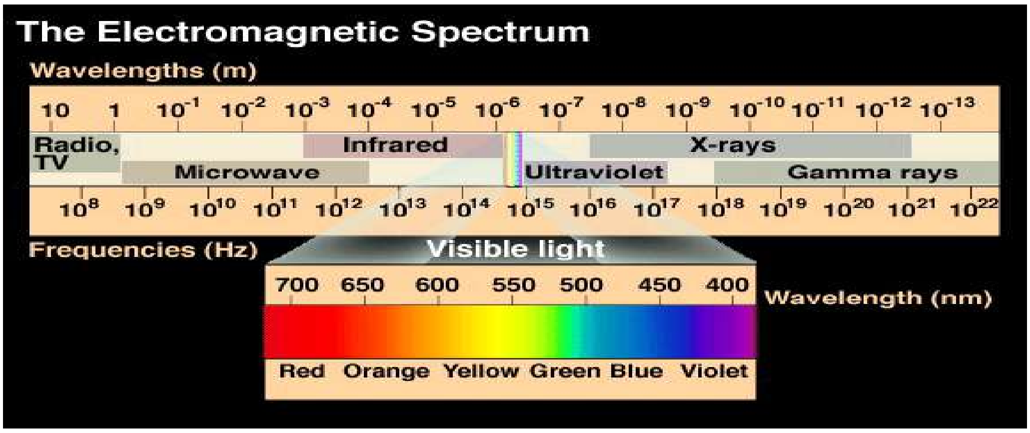
Wien's constant is a physical constant that defines the relationship between the thermodynamic temperature of a black body (an object that radiates electromagnetic energy perfectly) and the wavelength at which the intensity of the radiation is the greatest. The constant is denoted by the Greek lowercase letter sigma with a subscript w (σ_w). It is equal to approximately 2.898×10^{-3} meter-kelvin (0.2898 centimeter-kelvin).

The product of the thermodynamic temperature of a black body in kelvin s, and the wavelength of its peak energy output in meter s, is equal to Wien's constant. Thus, as a black body grows hotter, the wavelength of its peak energy output grows shorter



Electromagnetic Spectrum

The electromagnetic spectrum may be defined as the ordering of the radiation according to wavelength, frequency, or energy. The wavelength, denoted by A, is the distance between adjacent intensity maximum (for example) ofthe electromagnetic wave, and consequently, it may be expressedin any unit of length. Most commonly wavelength is expressed in meters (m) or centimeters (cm); microns or micrometers (11 or 11m = 10-4 cm); nanometers (nm = 10-7 cm); or Angstrom units (A = 10-8 cm).



Energy Interaction

Before radiation used for remote sensing reaches the Earth's surface it has to travel through some distance of the Earth's atmosphere. Particles and gases in the atmosphere can affect the incoming light and radiation. These effects are caused by the mechanisms of scattering and absorption.

- Atmosphere
  - Absorption (Ozone, Carbon Dioxide, Water Vapour)
  - Scattering (Rayleigh, Mie, Non-Selective)
- Earth's Surface
  - Vegetation
    - Absorption
  - Water
    - Transmission
  - Land
    - Reflection (Specular, Diffuse)

Absorption is the other main mechanism at work when electromagnetic radiation interacts with the atmosphere. In contrast to scattering, this phenomenon causes molecules in the atmosphere to absorb energy at various wavelengths. Ozone, carbon dioxide, and water vapour are the three main atmospheric constituents which absorb radiation.

Scattering occurs when particles or large gas molecules present in the atmosphere interact with and cause the electromagnetic radiation to be redirected from its original path. How much scattering takes place depends on several factors including the wavelength of the radiation, the abundance of particles or gases, and the distance the radiation travels through the atmosphere. There are three (3) types of scattering which take place.

Rayleigh scattering occurs when particles are very small compared to the wavelength of the radiation. These could be particles such as small specks of dust or nitrogen and oxygen molecules. Rayleigh scattering causes shorter wavelengths of energy to be scattered much more than longer wavelengths. Rayleigh scattering is the dominant scattering mechanism in the upper atmosphere. The fact that the sky appears "blue" during the day is because of this phenomenon. As sunlight passes through the atmosphere, the shorter wavelengths (i.e. blue) of the visible spectrum are scattered more than the other (longer) visible wavelengths. At sunrise and sunset the light has to travel farther through the atmosphere than at midday and the scattering of the shorter wavelengths is more complete; this leaves a greater proportion of the longer wavelengths to penetrate the atmosphere.

Mie scattering occurs when the particles are just about the same size as the wavelength of the radiation. Dust, pollen, smoke and water vapour are common causes of Mie scattering which tends to affect longer wavelengths than those affected by Rayleigh scattering. Mie scattering occurs mostly in the lower portions of the atmosphere where larger particles are more abundant, and dominates when cloud conditions are overcast.

The final scattering mechanism of importance is called**nonselective scattering**. This occurs when the particles are much larger than the wavelength of the radiation. Water droplets and large dust particles can cause this type of scattering. Nonselective scattering gets its name from the

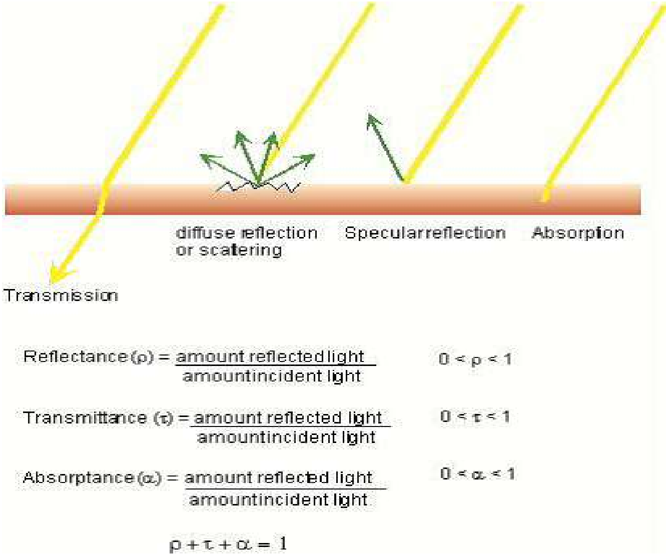
fact that all wavelengths are scattered about equally.
This type of scattering causes fog and clouds to appear white to our eyes because blue, green, and red light are all scattered in approximately equal quantities (blue+green+red light = white light).

EARTH SURFACE

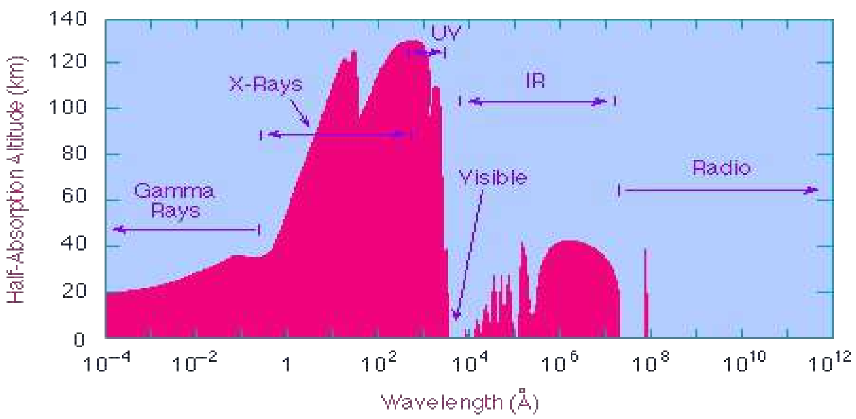
Radiation that is not absorbed or scattered in the atmosphere can reach and interact with the Earth's surface. There are three (3) forms of interaction that can take place when energy strikes, or is incident (I) upon the surface.
These are: absorption (A); transmission (T); and reflection (R). The total incident energy will interact with the surface in one or more of these three ways. The proportions of each will depend on the wavelength of the energy and the material and condition of the feature.
Absorption (A) occurs when radiation (energy) is absorbed into the target while transmission (T) occurs when radiation passes through a target.
Reflection (R) occurs when radiation "bounces" off the target and is redirected. In remote sensing, we are most interested in measuring the radiation reflected from targets. We refer to two types of reflection, which represent the two extreme ends of the way in which energy is reflected from a target.

Specular reflection and diffuse reflection.

When a surface is smooth, we get specular or mirror-like reflection where all (or almost all) of the energy is directed away from the surface in a single direction. Diffuse reflection occurs when the surface is rough and the energy is reflected almost uniformly in all directions. Most earth surface features lie somewhere between perfectly specular or perfectly diffuse reflectors. Whether a particular target reflects specularly or diffusely, or somewhere in between, depends on the surface roughness of the feature in comparison to the wavelength of the incoming radiation. If the wavelengths are much smaller than the surface variations or the particle sizes that make up the surface, diffuse reflection will dominate. For example, fine grained sand would appear fairly smooth to long wavelength microwaves but would appear quite rough to the visible wavelengths.



Atmospheric Windows: The general atmospheric transmittance across the whole spectrum of wavelengths is shown in figure below. The atmosphere selectively transmits energy of certain wavelengths. The spectral bands for which the atmosphere is relatively transparent are known as atmospheric windows. Atmospheric windows are present in the visible part (.4 μm - .76 μm) and the infrared regions of the EM spectrum. In the visible part transmission is mainly effected by ozone absorption and by molecular scattering. The atmosphere is transparent again beyond about λ= 1mm, the region used for microwave remote sensing.



What is a spectral signature

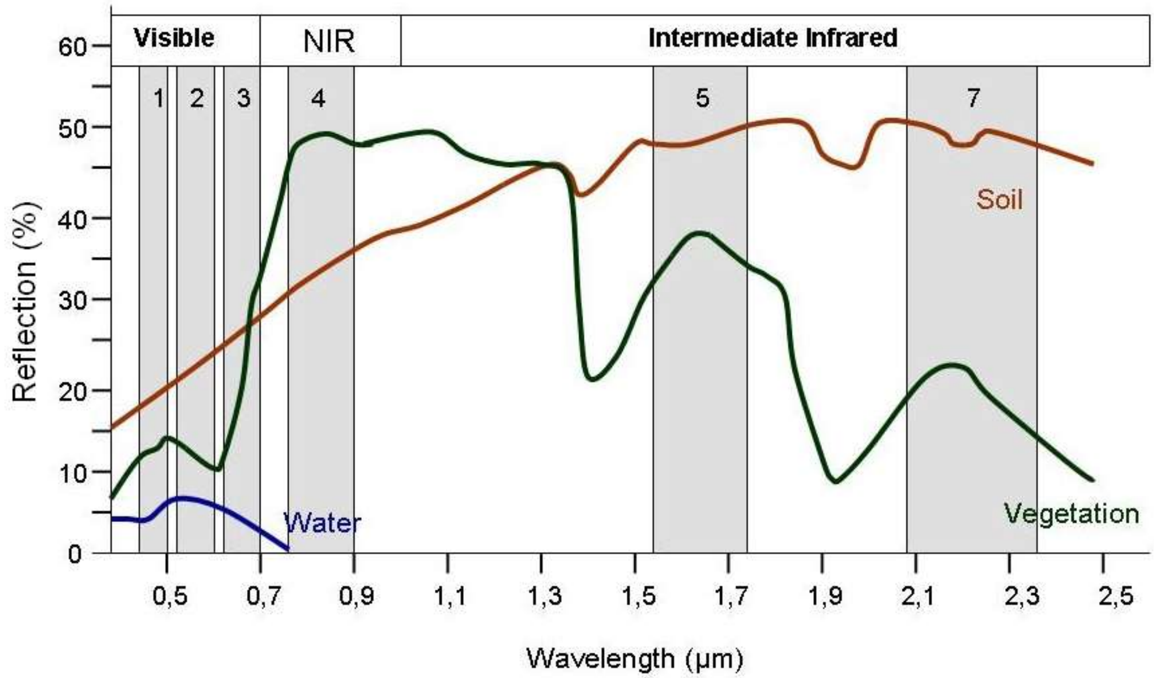
Features on the Earth reflect, absorb, transmit, and emit electromagnetic energy from the sun. Special digital sensors have been developed to measure all types of electromagnetic energy as it interacts with objects in all of the ways listed above. The ability of sensors to measure these interactions allows us to use remote sensing to measure features and changes on the Earth and in our atmosphere. A measurement of energy commonly used in remote sensing of the Earth is reflected energy (e.g., visible light, near-infrared, etc.) coming from land and water surfaces. The amount of energy reflected from these surfaces is usually expressed as a percentage of the amount of energy striking the objects. Reflectance is 100% if all of the light striking and object bounces off and is detected by the sensor. If none of the light returns from the surface, reflectance is said to be 0%. In most cases, the reflectance value of each object for each area of the electromagnetic spectrum is somewhere between these two extremes. Across any range of wavelengths, the percent reflectance values for landscape features such as water, sand, roads, forests, etc. can be plotted and compared. Such plots are called "spectral response curves" or "spectral signatures." Differences among spectral signatures are used to help classify remotely sensed images into classes of landscape features since the spectral signatures of like features have similar shapes. The figure below shows differences in the spectral response curves for healthy versus stressed sugar beet plants. spectral information recorded by a sensor, the more information that can be extracted from the spectral signatures. Hyperspectral sensors have much more detailed signatures than multispectral sensors and thus provide the ability to detect more subtle differences in aquatic and terrestrial features.

Spectral Reflectivity Reflectivity is the fraction of incident radiation reflected by a surface. The reflectance characteristics of Earth's surface features may be quantified by measuring the portion of incident energy that is reflected. This is measured as a function of wavelength (λ) and is called spectral reflectance ($r\lambda$)

Spectral reflectance of Water Clear water absorbs relatively little energy with wavelengths < 0.6 µm, resulting in high transmittance in the blue-green portion of the spectrum. As the turbidity of water changes, the reflectance changes dramatically. Increases in chlorophyll concentration tend to decrease reflectance in blue wavelengths and increase it in green wavelengths (can monitor algae)

Spectral reflectance of Vegetation. In the range from about 0.7 to 1.3 µm a plant leaf typically reflects 40 - 50% of the energy incident upon it primarily due to the internal structure of plant leaves. Because the internal structure of leaves are highly variable between plant species, reflectance measurements in this range often permit us to discriminate between species (even if they look the same in visible wavelengths). Many plant stresses alter the reflectance in this region, and sensors operating in this range are often used for vegetation stress detection

Spectral reflectance of Soil The factors that influence soil reflectance act over less specified spectral bands • Factors affecting soil reflectance are moisture content, soil texture (proportion of sand, silt and clay), surface roughness, presence of iron oxide and organic matter content. The presence of moisture in soil will decrease its reflectance - this effect is greatest in the water absorption bands at about 1.4, 1.9, 2.2 and 2.7 µm. Soil moisture content is strongly related to the soil texture



Spectral reflectance curves for Water, Soil, Vegetation

PLATFORMS AND SENSORS

Syllabus: Platforms and Sensors

Types of platforms-ground based, airborne, space based – orbit types, polar, Sun-synchronous and Geosynchronous –Optical and microwave sensors- Passive and Active sensors – resolution concept – spatial resolution, spectral resolution, radiometric, temporal resolution, Pay load description of important Earth Resources and Meteorological satellites, Resourcesat-I, Oceansat-I – Airborne and spaceborne TIR and microwave sensors

Learning Outcomes:

1. Understand the platform types used in remote sensing.
2. Identify the sensor appropriate for a particular application and concept of resolution.
3. Narrate the pay load description of important satellites.

PLATFORMS AND SENSORS

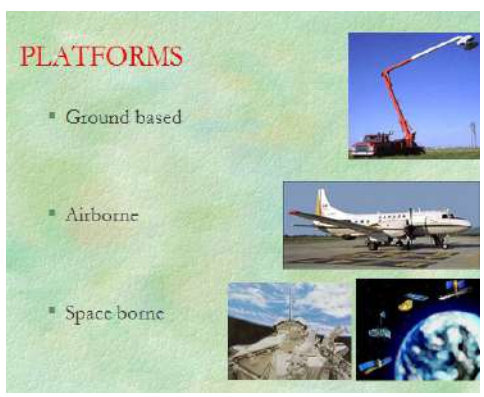
Platforms

A platform is the vehicle or carrier for remote sensors on which they are mounted. Platforms are used to house sensors which obtain data for remote sensing purposes, and are classified according to their heights and events to be monitored.

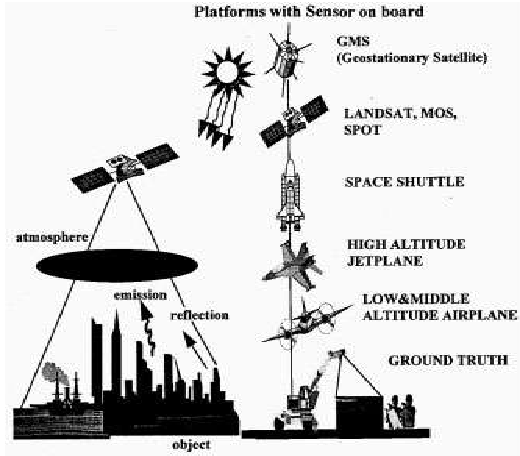
Permanent ground platforms are typically used for monitoring atmospheric phenomenon although they are also used for long-term monitoring of terrestrial features. Towers and cranes are often used to support research projects where a reasonably stable, long-term platform is necessary. Towers can be built on site and can be tall enough to project through a forest canopy so that a range of measurements can be taken from the forest floor, through the canopy and from above the canopy.

Platforms used for remote sensingare classified into three categories.

- Ground-based
- Airborne
- Space-borne



PLATFORMS AND SENSORS



Ground-Based Platforms

Ground-based sensors are often used to record detailed information about the surface which is compared with information collected from aircraft or satellite sensors. In some cases, this can be used to better characterize the target which is being imaged by these other sensors, making it possible to better understand the information in the imagery.

Ground based sensors may be placed on a ladder, scaffolding, tall building, cherry-picker, crane, etc.

Mobile Hydraulic Platforms

· Carried on vehicles

· Extendable to a height of 15m above the surface.

· At the top of the platform there are:

- Spectral reflectance meters

PLATFORMS AND SENSORS

- Photographic systems
- IR or Microwave scanners

Linked to data loggers in the vans.

Limitation: vehicles limited to roads, and the range is confined to small area along or around the road.

Portable Masts :

Used to support cameras and scanners

Limitation: could be unstable in windy conditions.

Towers :

Can be dismantled and moved from one place to another.

Offer greater rigidity than masts but are less mobile and require more time to erect.

PLATFORMS AND SENSORS

Airborne platforms

Airborne platforms are primarily stable wing aircraft, although helicopters are occasionally used. Aircraft are often used to collect very detailed images and facilitate the collection of data over virtually any portion of the Earth's surface at any time.

Balloon based missions and measurements :

High flying balloons provide an important tool for probing the atmosphere. Such balloonlaunches form an essential part of high altitude atmospheric research.There are three major advantages of the balloon programme including the following:

1. The balloon has extensive altitude range they can cover. They provide a unique way ofcovering a broad range of altitudes for in-situ or remote sensing measurements in thestratosphere. Of particular interest is the 22-40 km height range

2. The balloon instruments provide the opportunity for additional, correlative data forsatellite based measurements, including both validation ("atmospheric truth") andcomplementary data (for example, measurement of species not measured from the spacebased instrument).

3. Balloon based platforms constitute an important and inexpensive venue for testinginstruments under development. These can be either potential instruments for unmannedaerial vehicles (UAV) or, in some cases, for satellite based remote sensing instruments.

Here are some remote sensing instruments used in air borne remote sensing.

Radiosonde :

Radiosonde, an airborne instrument used for measuring pressure, temperature and relativehumidity in the upper air.

The instrument is carried aloft by a meteorological balloon inflated with hydrogen.

The radiosonde has a built-in high frequency transmitter that transmits data from theradiosonde meter and recorded on the ground by a specially designed radiosonde receiver.

PLATFORMS AND SENSORS

Rawinsonde

The rawinsonde is an electronic device used for measuring wind velocity, pressure,temperature and humidity aloft. It is also attached to a balloon and as it rises through the atmosphere, it measures the velocity of wind at respective altitudes.

Wind Finding Radar :

It determines the speed and direction of winds aloft by means of radar echoes. A radartarget is attached to a balloon and it is this target that is tracked by ground radar. Bearing and time of interval of the echoes is evaluated by a receiver.

Aircrafts are commonly used as remote-sensing for obtaining Aerial Photographs. In India, four types of aircrafts are being used for remote sensing operations.

DAKOTA : The ceiling height is 5.6 to 6.2 km and minimum speed is 240 km./hr.
AVRO : Ceiling height is 7.5 km and minimum speed is 600 km./hr.
CESSNA : Ceiling height is 9 km. and minimum speed is 350 km./hr.
CANBERRA: Ceiling height is 40 km. and minimum speed is 560 km./hr.

- In airborne remote sensing, downward or sideward looking sensors are mounted on anaircraft to obtain images of the earth's surface.
- An advantage of airborne remote sensing, compared to satellite remote sensing, is thecapability of offering very high spatial resolutionimages (20 cm or less).
- The disadvantages are low coverage area and high cost per unit area of ground coverage.
- It is not cost-effective to map a large area using an airborne remote sensing system.
- Airborne remote sensing missions are often carried out as one-time operations, whereasearth observation satellites offer the possibility of continuous monitoring of the earth.
- Analog aerial photography, videography, and digital photography are commonly used inairborne remote sensing.
- Synthetic Aperture Radarimaging is also carried out onairborne platforms.
- Analog photography is capable of providing high spatial resolution.

PLATFORMS AND SENSORS

- Digital photography permits real-time transmission of the remotely sensed data to aground station for immediate analysis.

Aircraft have several useful advantages as platform for remote sensing systems.

- Aircraft can fly at relatively low altitudes thus allowing for sub-meter sensor spatialresolution.
- Aircraft can easily change their schedule to avoid weather problems such as clouds, whichmay block a passive sensor's view of the ground.
- Last minute timing changes can be made to adjust for illumination from the sun, thelocation of the area to be visited and additional revisits to that location.
- Sensor maintenance, repair and configuration changes are easily made to aircraftplatforms.
- Aircraft flight paths know no boundaries except political boundaries.

Disadvantages of aircraft as platforms in remote sensing.

- Getting permission to intrude into foreign airspace can be a lengthy and frustratingprocess.
- The low altitude flown by aircraft narrows the field of view to the sensor requiring manypasses to cover a large area on the ground.
- The turnaround time it takes to get the data to the user is delayed due to the necessity ofreturning the aircraft to the airport before transferring the raw image data to the dataprovider's facility for preprocessing.

Space borne platforms :

In space, remote sensing is sometimes conducted from the space shuttle or, more commonly, from satellites. Satellites are objects which revolve around another object - in this case, the Earth.

For example, the moon is a natural satellite, whereas man-made satellites include those platforms launched for remote sensing, communication, and telemetry (location and navigation) purposes.

PLATFORMS AND SENSORS

Space borne platforms include the following

Rockets, Satellites and space shuttles. Space borne platforms range from 100 to 36000 kmabove the earth's surface.

Space shuttle: 250-300 km

Space stations: 300-400 km

Low level satellites: 700-1500 km

High level satellites: About 36000 km

Space borne remote sensing provides the following advantages:

- Large area coverage;
- Frequent and repetitive coverage of an area of interest;
- Quantitative measurement of ground features using radiometrically calibrated sensors;
- Semi automated computerised processing and analysis;
- Relatively lower cost per unit area of coverage.

Salient feature of some important satellite platforms.

Features	**Landsat1,2,3**	**Landsat 4,5**	**SPOT**	**IRS-IA**	**IRS-IC**
Nature	Sun Syn	Sun Syn	Sun Syn	Sun Syn	Sun Syn
Altitude (km)	919	705	832	904	817
Orbital period (minutes)	103.3	99	101	103.2	101.35
inclination (degrees	99	98.2	98.7	99	98.69
Temporal resolution (days)	18	16	26	22	24
Revolutions	251	233	369	307	341
Equatorial crossing (AM)	09.30	09.30	10.30	10.00	10.30
Sensors	RBV, MSS	MSS, TM	HRV	LISS-I, LISS-II	LISS-III, PAN, WIFS

PLATFORMS AND SENSORS

PLATFORMS AND SENSORS

Sensors

The sensoris a device used to acquire a photograph or an image. It will 'sense' and measures the amounts of radiated energy reflected from an object and record it. The types of sensors used are capable of capturing radiation from many different parts of the electromagnetic spectrum which are not visible to the human eye. Although the camera is a type of sensor, the word 'sensor' is normally used for the device used to acquire images in remote sensing.

The amount and range of the radiation that the sensor is capable of sensing, is specific to each type of sensor. When it is required to study and monitor some phenomena, first, scientists have to decide what type of pictures need to be taken of the area. The sensor used for the purpose will depend on the type of picture that needs to be taken

A camera is a sensor used in aerial photography. Black and white aerial photographs are normally taken with either panchromatic film or infrared-sensitive film.

Sensors are broadly of two types:

1. Optical sensor

Optical sensors observe visible lights and infrared rays (near infrared, intermediate infrared, thermal infrared). There are two kinds of observation methods using optical sensors:

a. Visible / Near IR

i. Method to acquire visible light and near infrared rays of sunlight *reflected*by objects on the ground.

ii. This method cannot be observed during night and cloud cover

b.Thermal Infrared

i. method to acquire thermal infrared rays, which is *radiat**ed***from land surface heated by sunlight.

ii. This method can be observed during night and cloud cover

PLATFORMS AND SENSORS

PLATFORMS AND SENSORS

2. Microwave sensor.

Microwave sensors receive microwaves, which are longer wavelength than visible light and infrared rays, and observation is not affected by day, night or weather.

There are two types of observation methods using microwave sensor:

a. Active

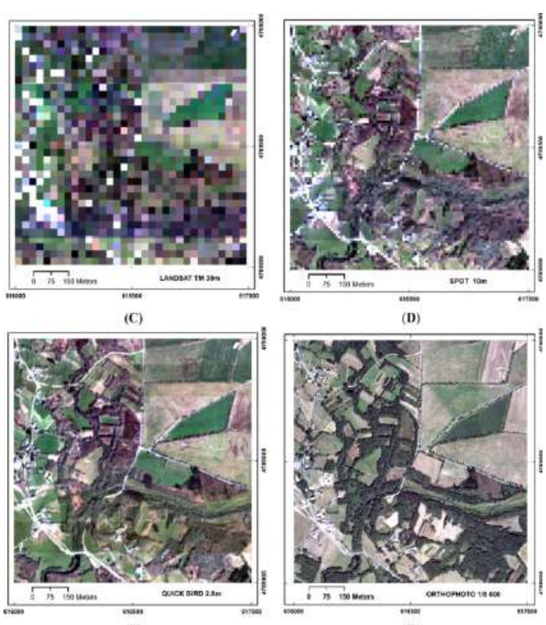
The sensor aboard earth observation satellite emits microwaves and observes microwaves reflected by land surface. It is suitable to observe mountains and valleys.

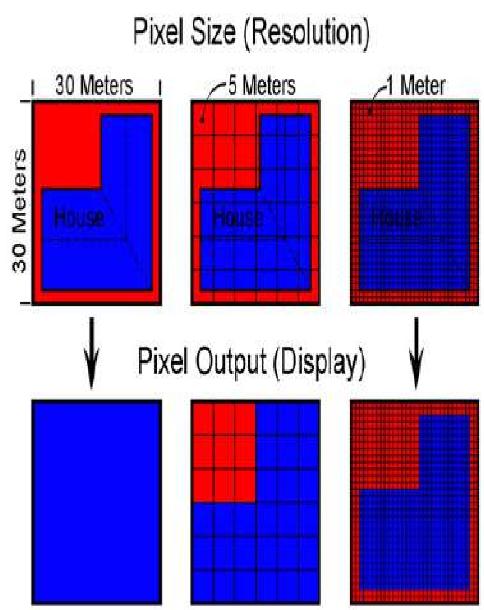
b. Passive

This type observes microwaves naturally radiated from land surface. It is suitable to observe sea surface temperature, snow accumulation, thickness of ice.

Resolution concept

Spatial resolution: It is a measure of the smallest angular or linear separation between two objects that can be resolved by the sensor. The greater the sensor's resolution, the greater the data volume and smaller the area covered. In fact, the area coverage and resolution are inter-dependant and these factors determine the scale of the imagery.



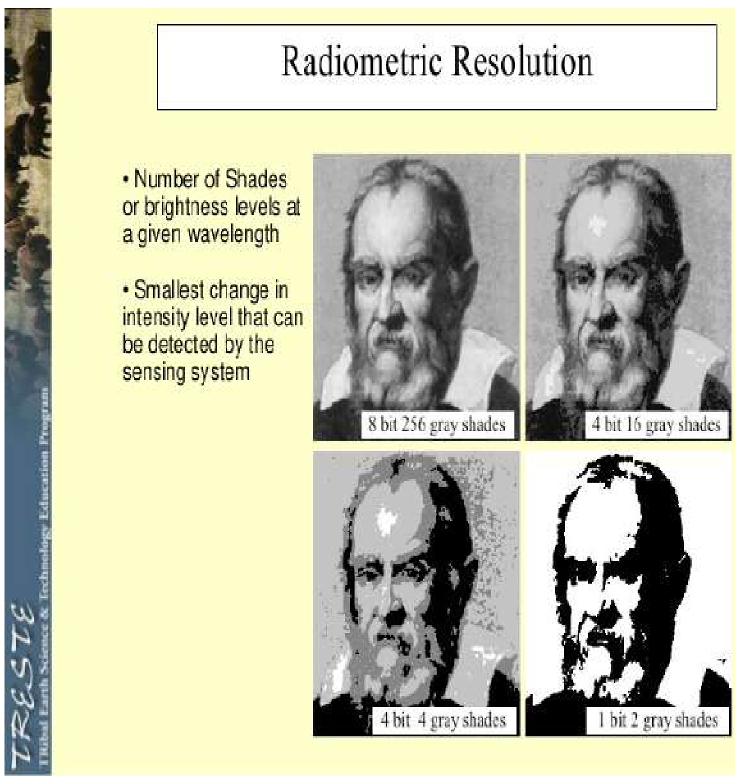


PLATFORMS AND SENSORS

Spectral resolution: It refers to the dimension and number of specific wavelength intervals in the electromagnetic spectrum to which a sensor is sensitive. Narrow bandwidths in certain regions of the electromagnetic spectrum allow the discrimination of various features more easily.

Temporal resolution: It refers to how often a given sensor obtains imagery of a particular area. Ideally, the sensor obtains data repetitively to capture unique discriminating characteristics of the phenomena of interest.

Radiometric sensitivity: It is the capability to differentiate the spectral reflectance/ emittance from various targets. This depends on the number of quantisation levels within the spectral band. In other words, the number of bits of digital data in the spectral band will decide the sensitivity of the sensor.



PLATFORMS AND SENSORS

Pay load description of Earth Resources and Meteorological satellites

CARTOSAT-1:

It will have a cutting-edge technology in terms of sensor systems and will provide state-of-art capabilities for cartographic applications. The satellite will have only a PAN camera with 2.5 m resolution and 30 km swath and Fore-Aft stereo capability. The 2.5 m resolution data will cater to the specific needs of cartography and terrain modeling applications.

RESOURCESAT-1:

Launched on 17th October, 2003, it is designed mainly for resources applications and having 3-band multi-spectral LISS-4 camera with a spatial resolution 5.8m and a swath of around 24 km with across – track steerability for selected area monitoring. An improved version of LISS-III, with 4 bands (green, red, near—IR and SWIR), all at 23.5 meters resolution and 140 km swath will also provide the much essential continuity to LISS-III. These payloads will provide enhanced data for vegetation applications and will allow multiple crop discrimination; species level discrimination and so on. Together with an advanced wide-field sensor, WiFS with ~ 60 m resolution and ~ 740 km swath, the payloads will aid greatly for crop and vegetation applications and integrated land and water applications. The data will also be useful for high accuracy resources management applications, where the emphasis is on multi crop mapping studies, vegetation species identification and utilities mapping.

INSAT-3D

INSAT-3D is an advanced weather satellite of India configured with improved Imaging System and Atmospheric Sounder. INSAT-3D is designed for enhanced meteorological observations, monitoring of land and ocean surfaces, generating vertical profile of the atmosphere in terms of temperature and humidity for weather forecasting and disaster warning.

It carries four payloads -

- 6 channel multi-spectral Imager
- 19 channel Sounder
- Data Relay Transponder (DRT)
- Search and Rescue Transponder

PLATFORMS AND SENSORS

The payloads of INSAT-3Dprovides continuity and further augment the capability to provide various meteorological as well as search and rescue services.

Orbits

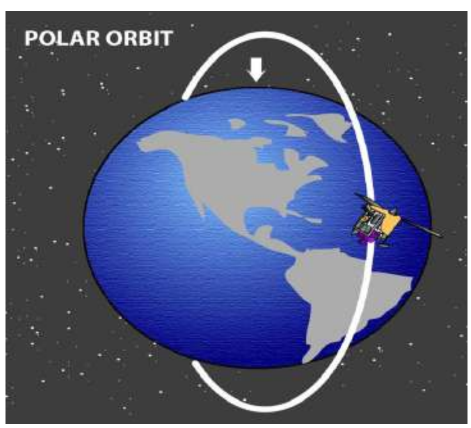
An orbit is a curved path of a celestial object around another celestial object due to the force of gravity. Orbits are everywhere in our universe. The Moon orbits the Earth, and the Earth orbits the Sun, and the Sun orbits around the center of the galaxy.

Types of Orbits

1. Polar
2. Sun Synchronous
3. Geosynchronous

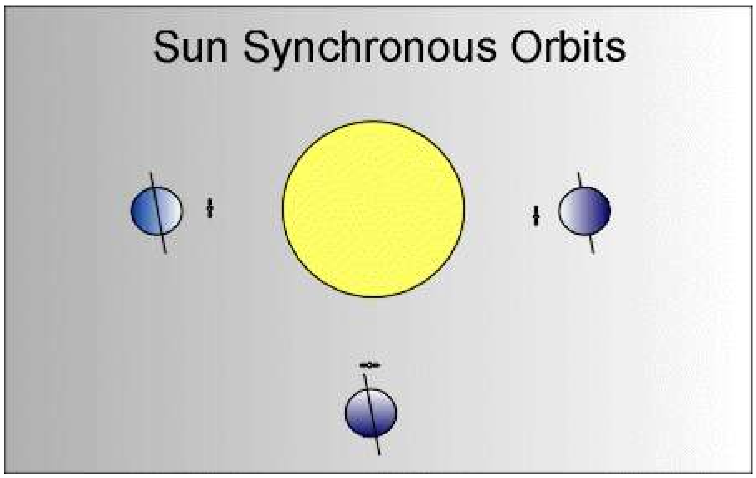
1. Polar Orbits

The more correct term would be near polar orbits. These orbits have an inclination near 90 degrees. This allows the satellite to see virtually every part of the Earth as the Earth rotates underneath it. It takes approximately 90 minutes for the satellite to complete one orbit. These satellites have many uses such as measuring ozone concentrations in the stratosphere or measuring temperatures in the atmosphere.



2. Sun Synchronous Orbits

These orbits allows a satellite to pass over a section of the Earth at the same time of day. Since there are 365 days in a year and 360 degrees in a circle, it means that the satellite has to shift its orbit by approximately one degree per day. These satellites orbit at an altitude between 700 to 800 km. These satellites use the fact since the Earth is not perfectly round (the Earth bulges in the center, the bulge near the equator will cause additional gravitational forces to act on the satellite. This causes the satellite's orbit to either proceed or recede. These orbits are used for satellites that need a constant amount of sunlight. Satellites that take pictures of the Earth would work best with bright sunlight, while satellites that measure longwave radiation would work best in complete darkness.

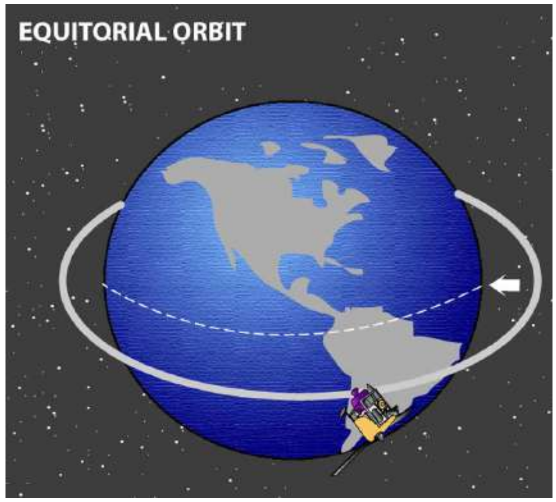


3. Geosynchronous Orbits

Also known as geostationary orbits, satellites in these orbits circle the Earth at the same rate as the Earth spins. The Earth actually takes 23 hours, 56 minutes, and 4.09 seconds to make one full revolution. So based on Kepler's Laws of Planetary Motion, this would put the satellite at approximately 35,790 km above the Earth. The satellites are located near the equator since at this latitude, there is a constant force of gravity from all directions. At other latitudes, the bulge at the center of the Earth would pull on the satellite.

Geosynchronous orbits allow the satellite to observe almost a full hemisphere of the Earth. These satellites are used to study large scale phenomenon such as hurricanes, or cyclones. These orbits are also used for communication satellites. The disadvantage of this type of orbit is that since these satellites are very far away, they have poor resolution. The other disadvantage is that these satellites have trouble monitoring activities near the poles. See the picture below.

PLATFORMS AND SENSORS



UNIT – III

IMAGE INTERPRETATION AND ANALYSIS

Types of Data Products- raw data, partially processed data, orthorectified data,special data products,– types of image interpretation– basic elements of image interpretation - visual interpretation keys – Digital Image Processing – Preprocessing – image enhancement techniques – multispectral image classification – Supervised- minimum distance/centroid classifier, maximum likelihood and unsupervised-open pass cluster, sequential cluster, statistical cluster, k-means cluster.

Learning Outcomes:

1. Understand the data types in remote sensing.
2. Discuss the elements of image interpretations.
3. Discuss the types of classification.

Types of Remote sensing Data products :

1. Raw data: No processing is done and the relevant scene scene is extracted and put on media.
2. Partially processed data: Radiometrically corrected but not geometrically or vice versa.
3. Processed or standard data: these data are both radiometrically and geometrically corrected.
4. Geocoded data: these products are north oriented and compatible to the survey of indiatoposheets.
5. Ortho-rectified: orthorectified images are geometrically corrected products with corrections for displacement caused by terrain and relief.
6. Special data products: besides the above data products NRSC and other agencies develop some custom data products for their requirements.

BASIC PRINCIPLES OF IMAGE INTERPRETATION :

Images and their interpretability

- An image taken from the air or space is a pictorial presentation of the pattern of a landscape.
- The pattern is composed of indicators of objects and events that relate to the physical, biological and cultural components of the landscape.

- Similar conditions, in similar circumstances and surroundings, reflect similar patterns, and unlike conditions reflect unlike patterns.
- The type and amount of information that can be extracted is proportional to the knowledge, skill and experience of the analyst, the methods used for interpretation and the analyst's awareness of any limitations.

Factors Governing the Quality of an image

In addition to the inherent characteristics of an object itself, the following factors influence image quality:

- Sensor characteristics (film types, digital systems),
- Season of the year and time of day
- Atmospheric effects,
- Resolution of the imaging system and Scale
- Image motion
- Stereoscopic parallax

Factors Governing Interpretability

- Experience of the interpreter (Field knowledge, Common Sense, Subject, Similar Works)
- Equipment and technique of interpretation
- Interpretation keys, guides, manuals and other aids.

Visibility of Objects

The objects on aerial photographs or imagery are represented in the form of photo images in tones of grey in B/W photography and in colour/false colour photography in different colours/hues. This visibility of objects in the images varies due to

- The inherent characteristics of the objects
- The quality of the aerial photographs.

Elements of Image Interpretation	
Primary Elements	Black and White Tone
	Color
	Stereoscopic Parallax
Spatial Arrangement of Tone and Color	Size
	Shape
	Texture
	Pattern
Based on Analysis of Primary Elements	Height
	Shadow
Contextual Elements	Site
	Association

Basic elements of image interpretation:

(i) Tone: Ground objects of different color reflect the incident radiation differently depending

upon the incident wave length, physical and chemical constituents of the objects. The imagery as recorded in remote sensing is in different shades or tones.

(ii) Texture: Texture is an expression of roughness or smoothness as exhibited by the imagery. It is the rate of change of tonal values. Texture can qualitatively be expressed as course, medium and fine. The texture is a combination of several image characteristics such as tone, shadow size, shape and pattern etc., and is produced by a mixture of features too small to be seen individually because the texture by definition is the frequency of tonal changes.

(iii) Association: The relation of a particular feature to its surroundings is an important key to interpretation. Some times a single featured by itself may not be distinctive enough to permit

its identification.

(iv) Shape: Some ground features have typical shapes due to the structure or topography. For example air fields and football stadium easily can be interpreted because of their finite ground shapes and geometry whereas volcanic covers, sand, river terraces, cliffs, gullies can be identified because of their characteristic shape controlled by geology and topography.

(v) Size: The size of an image also helps for its identification whether it is relative or absolute. Sometimes the measurements of height (as by using parallax bar) also gives clues to the nature of the object.

(vi) Shadows: shadows cast by objects are sometimes important clues to their identification and interpretation. For example, shadow of a suspension bridge can easily be discriminated from that of cantilever bridge. Similarly circular shadows are indicative of coniferous trees.

(vii) Size factor or topographic location: Relative elevation or specific location of objectscan be helpful to identify certain features. For example, sudden appearance or disappearance of vegetation is a good clue to the underlying soil type or drainage conditions.

(viii) Pattern: Pattern is the orderly special arrangement of geological topographic or vegetation features. This special arrangement may be two-dimensional or 3-dimensional.

IMAGE PROCESSING

Image Processing and Analysis can be defined as the "act of examining images for thepurpose of identifying objects and judging their significance.

Digital Data: Remote sensing images are recorded in digital forms. A digital image is anarray of numbers depicting spatial distribution of a certain field parameters based onreflectivity of EM radiation of the ground feature.

Digital Image Processing is a collection of techniques for the manipulation of digital imagesby computers. The raw data received from the imaging sensors on the satellite platforms
contains flaws and deficiencies. To overcome these flaws and deficiencies in order to get theoriginality of the data, it needs to undergo several steps of processing. This will vary fromimage to image depending on the type of image format, initial condition of the imageand theinformation of interest and the composition of the image scene. Digital Image Processingundergoes three general steps:

_ Pre-processing
_ Display and enhancement
_ Information extraction

Preprocessing

Remotely sensed raw data, received from imaging sensor mounted on satellite platformsgenerally contain flaws and deficiences. The correction of deficiences and removal of flaws present in the data through some methods are termed as pre-processing methods. This correction model involves the initial processing of raw image data to correct geometric distortions, to calibrate the data radiometrically and to eliminate the noise present in the data. All pre-processing methods are considered under three heads, namely, (i) geometric correction methods, (ii) radiometric correction methods, and (iii) atmospheric correction methods. (iv) Feature Extraction

1. Geometric Correction Methods

Remotely sensed images are not maps (Mather, 2000). Frequently information extracted from remotely sensed images is Integrated with map data in a geographical information system. The transformation of a remotely sensed image into a map with a scale and projection properties is called geometric correction

2. Radiometric Correction Methods

The primary function of remote sensing data quality evaluation is to monitor the performance of the sensors. The performance of the sensors is continuously monitored by applying radiometric correction models on digital Image data sets. The radiance measured by any given system over a given object is influenced by factors, such as, changes in scene illumination, atmospheric conditions, viewing geometry and instrument response characteristics (Lillesand and Kiefer, 2000). One of the most important radiometric data processing activity involved in many quantitative applications of digital image data is conversion of digital numbers to absolute physical values; namely, radiance and reflectance.

3 Atmospheric Correction Methods

According to Rayleigh scattering, the effect of scattering is inversely proportional to the fourth power of wavelength of energy, that is, scattering is more in the lower wavelength (visible) than in the higher wavelength (infrared band). Further scattering effect increases the signal value (bias). Let us consider a case in which the energy in various wavelengths is interacting with the earth surface features, for example, water body. Assume that the sky is very clear and there is no atmospheric scattering or haze. If the sky is clear with no scattering, then the radiance reflected from the earth surface feature in any of the region of the electromagnetic spectrum should be the same. This is the ideal case. In reality, because of the presence of haze, fog, or atmospheric scattering, there always exists some kind of unwanted signal value called bias.

4. Feature Extraction

Feature Extraction does not mean geographical features visible on the image but rather"statistical" characteristics of image data like individual bands or combination of bandvalues that carry information concerning systematic variation within the scene. Thus in amultispectral data it helps in portraying the necessity elements of the image. It alsoreduces the number of spectral bands that has to be analyzed. After the feature extractionis complete the analyst can work with the desired channels or bands, but in turn theindividual bandwidths are more potent for information. Finally such a preprocessingincreases the speed and reduces the cost of analysis.

IMAGE ENHANCEMENT TECHNIQUES

Image Enhancement techniques are instigated for making satellite imageries moreinformative and helping to achieve the goal of image interpretation. The term enhancement isused to mean the alteration of the appearance of an image in such a way that the informationcontained in that image is more readily interpreted visually in terms of a particular need. Theimage enhancement techniques are applied either to single-band images or separately to theindividual bands of a multiband image set.

Image reduction

The number of columns and rows in a remotely sensed image scene generally exceeds the screen resolution of the computer screen, so in order to reduce the image every nth row and nth column of the image are systematically selected and displayed.

Image Magnification

Digital image magnification is oftenlyrefered to a zoom in. This technique is most comonly employed for two purposes.

1. To improve the display scale of the image for enhanced visual interpretation.
2. To match the display scale of another image

Color composition

Color composition is the assignment of three primary colors; red (R), green (G) and blue (B) to three selected bands from multispectral bands usually available for satellite remote sensing image data.

Transect extraction

Users of remotely sensed imagery frequently extract brightness values between points on an image. a transect is a straight line between any two user-specifies points with in an image

Contrast Enhancement

The operating or dynamic, ranges of remote sensors are often designed with a variety ofeventual data applications. For example for any particular area that is being imaged it isunlikely that the full dynamic range of sensor will be used and the corresponding image isdull and lacking in contrast or over bright. Landsat TM images can end up being used tostudy deserts, ice sheets, oceans, forests etc., requiring relatively low gain sensors to copewith the widely varying radiances upwelling from dark, bright , hot and cold targets.

Consequently, it is unlikely that the full radiometric range of brand is utilised in an image of aparticular area. The result is an image lacking in contrast - but by remapping the DNdistribution to the full display capabilities of an image processing system, we can recover abeautiful image. Contrast Stretching can be displayed in three catagories

Linear Contrast Stretch

This technique involves the translation of the image pixel values from the observed rangeDNmin to DNmax to the full range of the display device(generally 0-255, which is the rangeof values representable in an 8bit display devices)This technique can be applied to a singleband, grey-scale image, where the image data are mapped to the display via all three colorsLUTs.

It is not necessary to stretch between DNmax and DNmin - Inflection points for a linearcontrast stretch from the 5th and 95th percentiles, or ± 2 standard deviations from the mean(for instance) of the histogram, or to cover the class of land cover of interest (e.g. water atexpense of land or vice versa). It is also straightforward to have more than two inflectionpoints in a linear stretch, yielding a piecewise linear stretch.

Histogram Equalisation

The underlying principle of histogram equalisation is straightforward and simple, it isassumed that each level in the displayed image should contain an approximately equalnumber of pixel values, so that the histogram of these displayed values is almost uniform(though not all 256 classes are necessarily occupied). The objective of the histogramequalisation is to spread the range of pixel values present in the input image over the fullrange of the display device.

Density Slicing

Density Slicing is the mapping of a range of contiguous grey levels of a single band image to a point in the RGB color cube. The DNs of a given band are "sliced" into distinct classes. For example, for band 4 of a TM 8 bit image, we might divide the 0-255 continuous range into discrete intervals of 0-63, 64-127, 128-191 and 192-255. These four classes are displayed as four different grey levels. This kind of density slicing is often used in displaying temperature maps.

Filtering

Filtering encompasses another set of digital processing functions, which are used to enhance the appearance of an image.

Spatial filters are designed to highlight or suppress specific features in an image based on their spatial frequency.

Image Classification

Image classification is a procedure to automatically categorize all pixels in an image of a terrain into land cover classes. Normally, multispectral data are used to perform the classification of the spectral pattern present within the data for each pixel

Digital Image Processing is used as the numerical basis for categorization. This concept is dealt under the broad subject, namely, Pattern Recognition. Spectral pattern recognition refers to the family of classification procedures that utilises this pixel-by-pixel spectral information as the basis for automated land cover classification. Spatial pattern recognition involves the categorization of image pixels on the basis of the spatial relationship with pixels surrounding them. Image classification techniques are grouped into two types, namely supervised and unsupervised. The classification process may also include features, such as, land surface elevation and the soil type that are not derived from the image. A pattern is thus a set of measurements on the chosen features for the individual to be classified. The classification process may therefore be considered a form of pattern recognition, that is, the identification of the pattern associated with each pixel position in an image in terms of the characteristics of the objects or on the earth's surface.

1 Supervised Classification

A supervised classification algorithm requires a training sample for each class, that is, a collection of data points known to have come from the class of interest. The classification is thus based on how "close" a point to be classified is to each training sample. We shall not attempt to define the word "close" other than to say that both geometric and statistical distance measures are used in practical pattern recognition algorithms. The training samples are representative of the known classes of interest to the analyst. Classification methods that relay on use of training patterns are called supervised classification methods. The three basic steps involved in a typical supervised classification procedure are as follows :

(i) Training stage: The analyst identifies representative training areas and develops numerical descriptions of the spectral signatures of each land cover type of interest in the scene.

(ii) The classification stage: Each pixel in the image data set iscategorised into the land cover class it most closely resembles. If the pixel is insufficiently similar to any training data set it is usually labeled 'Unknown'.

(iii) The output stage: The results may be used in a number of different ways. Three typical forms of output products are thematic maps, tables and digital data files which become input data for GIS.

2 Unsupervised Classification

Unsupervised classification algorithms do not compare .points to be classified with training data. Rather, unsupervised algorithms examine a large number of unknown data vectors and divide them into classes based on properties inherent to the data themselves. The classes that result stem from differences observed in the data. In particular, use is made of the notion that data vectors within a class should be in some sense mutually close together in the measurement space, whereas data vectors in different classes should be comparatively well separated. If the components of the data vectors represent the responses in different spectral bands, the resulting classes might be referred to as spectral classes, as opposed to information classes, which represent the ground cover types of interest to the analyst.

Supervised Classification

In this system each pixel is supervised for the categorization of the data by specifying to thecomputer algorithm, numerical descriptors of various class types. There are three basic stepsinvolved in typical supervised classification

Training Stage

The analyst identifies the training area and develops a numerical description of the spectralattributes of the class or land cover type. During the training stage the location, size, shapeand orientation of each pixel type for each class.

Classification Stage

Each pixel is categorised into landcover class to which it closely resembles. If the pixel is notsimilar to the training data, then it is labeled as unknown. Numerical mathematicalapproaches to the spectral pattern recognition have been classified into various categories.

1. Measurements on Scatter Diagram

Each pixel value is plotted on the graph as the scatter diagram indicating the category of theclass. In this case the 2-dimensional digital values attributed to each pixel is plottes on thegraph

2. Minimum Distance to Mean Classifier/Centroid Classifier

This is a simple classification strategies. First the mean vector for each category isdetermined from the average DN in each band for each class. An unknown pixel can then beclassified by computing the distance from its spectral position to each of the means andassigning it to the class with the closest mean. One limitation of this technique is that itoverlooks the different degrees of variation.

3. Parallelpiped Classifier

For each class the estimate of the maximum and minimum DN in each band is determine.Thenparallelpiped are constructeds o as to enclose the scatter in each theme. Then each pixelis tested to see if it falls inside any of the parallelpiped and has limitation

- A pixel may fall outside the parallelpiped and remained unclassified.
- Theme data are so strongly corrected such that a pixel vector that plots at some distance from the theme scatter may yet fall within the decision box and be classified erroneously.
- Sometimes parallelpiped may overlap in which case the decision becomes more
- complicated then boundary are slipped.

4. Gaussian Maximum Likelihood Classifier

This method determines the variance and covariance of each theme providing the probabilityfunction. This is then used to classify an unknown pixel by calculating for each class, theprobability that it lies in that class. The pixel is then assigned to the most likely class or if itsprobability value fail to reach any close defined threshold in any of the class, be labeled asunclassified. Reducing data dimensionally before handis an approach to speeding theprocess up.

Unsupervised Classification

This system of classification does not utilize training data as the basis of classification. Thisclassifier involves algorithms that examine the unknown pixels in the image and aggregatethem into a number of classes based on the natural groupings or cluster present in the image.The classes that result from this type of classification are spectral classes. Unsupervisedclassification is the identification, labeling and mapping of these natural classes. This methodis usually used when there is less information about the data before classification.

There are several mathematical strategies to represent the clusters of data in spectral space.

1. Sequential Clustering

In this method the pixels are analysed one at a time pixel by pixel and line by line. Thespectral distance between each analysed pixel and previously defined cluster means arecalculated. If the distance is greater than some threshold value, the pixel begins a new cluster

Otherwise it contributes to the nearest existing clusters in which case cluster mean isrecalculated. Clusters are merged if too many of them are formed by adjusting the thresholdvalue of the cluster means.

2. Statistical Clustering

It overlooks the spatial relationship between adjacent pixels. The algorithm uses 3x3windows in which all pixels have similar vector in space. The process has two stepsofTesting for homogeneity within the window of pixels under consideration.

o Cluster merging and deletion

Here the windows are moved one at time through the image avoiding the overlap. The meanand standard derivation is calculated for each band of the window. The smaller the standarddeviation for a given band the greater the homogenity of the window. These values are thencompared by the user specified parameter for delineating the upper and lower limit of thestandard deviation. If the window passes the homogeneity test it forms cluster. Clusters arecreated untill then number exceeds the user defined maximum number of clusters at whichpoint some are merged or deleted according to their weighting and spectral distances.

3. Iso Data Clustering (Iterative Self Organising Data Analysis Techniques)

Its repeatedly performs an entire classification and recalculates the statistics. The procedurebegins with a set of arbitrarily defined cluster means, usually located evenly through thespectral space. After each iteration new means are calculated and the process is repeated untilthere is some difference between iterations. This method produces good result for the datathat are not normally distributed and is also not biased by any section of the image.

4. RGB Clustering

It is quick method for 3 band, 8 bit data. The algorithm plots all pixels in spectral space andthen divides this space into 32 x 32 x 32 clusters. A cluster is required to have minimumnumber of pixels to become a class. RGB Clustering is not based to any part of the data.

UNIT-4

Geographical Information System

Introduction – Maps – Definitions, Map projections – types of map projections, map analysis GIS definition – basic components of GIS, Standard GIS software's, Data type, Spatial and non-spatial (attribute) data , Measurement scales, ata Base Management Systems(DBMS). Attribute data analysis – integrated data analysis

INTRODUCTION

Everything is related with Geography… Want to locate a city… Want to measure the length or area or height of a feature… want to know the distance between two places… Want to study the soil characteristics… Estimate the cost of construction… calculate the total vegetation or forest resources… etcetcetc ….

DEFINITIONS

GIS is defined as an information system that is used to **input, store, retrieve, manipulate, analyze and output** geographically referenced data or geospatial data, in order to support decision making for planning andmanagement of land use, natural resources, environment, transportation, urban facilities, health services so on.

MAP a diagrammatic representation of an area of land or sea showing physical features, cities, roads, etc.

Types of Maps

Topographic Map –a reference tool showing the outlines of selected natural and man-made features of earth –'topography' refers to the shape of the surface, represented by contours and/or shading –topographic maps also show roads and other prominent features •

Thematic Map –a tool to communicate geographical concepts such as the distribution of population densities, climate, movement of goods, landuse, etc

Map Scale

Map scale defines the amount of reduction of reality.

It is the ratio between distances on map and corresponding distances in the real world.

Scale is expressed in three primary ways

1.Verbal Scale

2.Representative fraction (RF)

3.Graphic scale (bar)

GIS is a set of tool that allows for the processing of spatial data into information.

TERMINOLOGIES

1.Geographical Entities

2.Attributes

3.Topology

4.Cognitive Models

CATEGORIES

2-D (x,y) ; 2.5-D : DTM [z = f(x,y)], 3-D : DEM (x,y,z) ; 4-D (Temporal GIS – based on time)

Components of a GIS

A GIS can be divided into five components: People, Data, Hardware, Software, and Procedures. All of these components need to be in balance for the system to be successful. No one part can run without the other.

People

The people are the component who actually makes the GIS work. They include a plethora of positions including GIS managers, database administrators, application specialists, systems analysts, and programmers. They are responsible for maintenance of the geographic database and provide technical support. People also need to be educated to make decisions on what type of system to use. People associated with a GIS can be categorized into: viewers, general users, and GIS specialists.

· Viewers are the public at large whose only need is to browse a geographic database for referential material. These constitute the largest class of users.

· General Users are people who use GIS to conducting business, performing professional services, and making decisions. They include facility managers, resource managers, planners, scientists, engineers, lawyers, business entrepreneurs, etc.

· GIS specialists are the people who make the GIS work. They include GIS managers, database administrators, application specialists, systems analysts, and programmers. They are responsible for the maintenance of the geographic database and the provision of technical support to the other two classes of users. (Lo, 2002)

Procedures

Procedures include how the data will be retrieved, input into the system, stored, managed, transformed, analyzed, and finally presented in a final output. The procedures are the steps taken to answer the question needs to be resolved. The ability of a GIS to perform spatial analysis and answer these questions is what differentiates this type of system from any other information systems.

The transformation processes includes such tasks as adjusting the coordinate system, setting a projection, correcting any digitized errors in a data set, and converting data from vector to raster or raster to vector. (Carver, 1998)

Hardware

Hardware consists of the technical equipment needed to run a GIS including a computer system with enough power to run the software, enough memory to store large amounts of data, and input and output devices such as scanners, digitizers, GPS data loggers, media disks, and printers. (Carver, 1998)

Software

There are many different GIS software packages available today. All packages must be capable of data input, storage, management, transformation, analysis, and output, but the appearance, methods, resources, and ease of use of the various systems may be very different. Today's software packages are capable of allowing both graphical and descriptive data to be stored in a single database, known as the object-relational model. Before this innovation, the geo-relational model was used. In this model, graphical and descriptive data sets were handled separately. The modern packages usually come with a set of tools that can be customized to the users needs (Lo, 2002).

Data

Perhaps the most time consuming and costly aspect of initiating a GIS is creating a database. There are several things to consider before acquiring geographic data. It is crucial to check the quality of the data before obtaining it. Errors in the data set can add many unpleasant and costly hours to implementing a GIS and the results and conclusions of the GIS analysis most likely will be wrong.

FUNDAMENTAL OPERATIONS OF GIS

(i) Reclassification operations transform the attribute information associated with a single map coverage.

(ii) Overlay operations involve the combination of two or more maps according to boolean conditions and may resultin the delineation of new boundaries.

(iii) Distance and connectivity measurement include both simple measure of interpoint distance and more complexoperations such as the construction of zones of increasing transport cost away from specified locations, and

(iv) Neighbourhood characterisation involves the values to a location both summary and mean measures of avariable, and include smoothing and enhancement filters. Sequences or such manipulation operations have become known as 'cartographic modelling'.

THEORETICAL FRAMEWORK OF GIS

Based on the way in which the data is transformed :

1. Traditional Cartographic process
2. Based on GIS operation.

GIS ARCHITECTURE

GIS WORKFLOW

5 elements

1. Data Acquisition

2. Preprocessing

a. Format Conversion

b. Data Reduction & Generalization

c. Error Detection & Editing

d. Merging the points into lines and lines into Polygons

e. Edge-matching & Tiling

f. Rectification & Registration

g. Interpolation

h. Interpretation

3. Data Management

a. Insert

b. Update

c. Delete

d. Retrieve

4. Manipulation & Analysis

a. Reclassification & Aggregation

b. Geometric Operations : Rotation, Scaling, Rectification

c. Controlled Determination

d. Data Structure Conversion

e. Spatial Operation of Connectivity &Neighbourhood Operations

f. Measurement of Distance & Direction

g. Statistical Analysis

h. Modeling

5. Product Generation (Final Outputs from GIS are created)

Georeferencing and Co-ordinate systems

Georeferencing: Assigning map coordinates and spatial location

All the elements in a map layer have a specific geographic location and extent that enables them to be located on or near the earth's surface. The ability to accurately describe geographic locations is critical in both mapping and GIS. This process is called georeferencing.

Describing the correct location and shape of features requires a framework for defining real-world locations. A geographic coordinate system is used to assign geographic locations to objects. A global coordinate system of latitude-longitude is one such framework. Another is a planar or Cartesian coordinate system derived from the global framework.

Maps represent locations on the earth's surface using grids, graticules, and tic marks labeled with various ground locations (both in measures of latitude-longitude and in projected coordinate systems (such as UTM meters). The geographic elements contained in various map layers are drawn in a specific order (on top of one another) for the given map extent.

GIS datasets contain coordinate locations within a global or Cartesian coordinate system to record geographic locations and shapes.

Latitude and longitude

One method for describing the position of a geographic location on the earth's surface is using spherical measures of latitude and longitude. They are measures of the angles (in degrees) from the center of the earth to a point on the earth's surface. This reference system is often referred to as a geographic coordinate system.

Latitude angles are measured in a north-south direction. The equator is at an angle of 0. Often, the northern hemisphere has positive measures of latitude and the southern hemisphere has negative measures of latitude. Longitude measures angles in an east-west direction. Longitude measures are traditionally based on the Prime Meridian, which is an imaginary line running from the North Pole through Greenwich, England to the South Pole. This angle is Longitude 0. West of the Prime Meridian is often recorded as negative Longitude and east is recorded as positive. For example, the location of Los Angeles, California is roughly Latitude "plus 33 degrees, 56 minutes" and Longitude "minus 118 degrees, 24 minutes."

Although longitude and latitude can locate exact positions on the surface of the globe, they are not uniform units of measure. Only along the equator does the distance represented by one degree of longitude approximate the distance represented by one degree of latitude. This is because the equator is the only parallel as large as a meridian. (Circles with the same radius as the spherical earth are called great circles. The equator and all meridians are great circles.)

Above and below the equator, the circles defining the parallels of latitude get gradually smaller until they become a single point at the North and South Poles where the meridians converge. As the meridians converge toward the poles, the distance represented by one degree of longitude decreases to zero. On the Clarke 1866 spheroid, one degree of longitude at the equator equals 111.321 km, while at 60° latitude, it is only 55.802 km. Since degrees of latitude and longitude don't have a standard length, you can't measure distances or areas accurately or display the data easily on a flat map or computer screen. Performing GIS analysis and mapping applications requires a more stable coordinate framework, which is provided by projected coordinate systems.

Map projections using Cartesian coordinates

Projected coordinate systems are any coordinate system designed for a flat surface, such as a printed map or a computer screen. 2D and 3D Cartesian coordinate systems provide the mechanism for describing the geographic location and shape of features using x and y values (and, as you will read later, by using columns and rows in rasters).

The Cartesian coordinate system uses two axes: one horizontal (x), representing east-west, and one vertical (y), representing north-south. The point at which the axes intersect is called the origin. Locations of geographic objects are defined relative to the origin, using the notation (x,y), where x refers to the distance along the horizontal axis, and y refers to the distance along the vertical axis. The origin is defined as (0,0).

In the illustration below, the notation (4, 3) records a point that is four units over in x and three units up in y from the origin.

3D coordinate systems

Increasingly, projected coordinate systems also use a Z value to measure elevation above or below mean sea level. In the illustration below, the notation (2, 3, 4) records a point that is two units over in x and three units in y from the origin and whose elevation is 4 units above the earth's surface (such as 4 meters above mean sea level).

Properties and distortion in map projections

Since the earth is spherical, a challenge faced by cartographers and GIS professionals is how to represent the real world using a flat or planar coordinate system. To understand their dilemma, consider how you would flatten half of a basketball; it can't be done without distorting its shape or creating areas of discontinuity. The process of flattening the earth is called projection, hence the term map projection.

A projected coordinate system is defined on a flat, two dimensional surface. Projected coordinates can be defined for both 2D (x,y) and 3D (x,y,z) in which the x,y measurements represent the location on the earth's surface and z would represent height above or below mean sea level.

Below are some examples of various methods for deriving planar map projections.

Unlike a geographic coordinate system, a projected coordinate system has constant lengths, angles, and areas across the two dimensions. However, all map projections representing the earth's surface as a flat map, create distortions in some aspect of distance, area, shape, or direction.

Users cope with these limitations by using map projections that fit their intended uses, geographic location, and extent. GIS software also can transform information between coordinate systems to support integration and critical workflows.

Many map projections are designed for specific purposes. One map projection might be used for preserving shape while another might be used for preserving the area (conformal versus equal area).

These properties—the map projection (along with Spheroid and Datum), become important parameters in the definition of the coordinate system for each GIS dataset and each map. By recording detailed descriptions of these properties for each GIS dataset, computers can re-project and transform the geographic locations of dataset elements on the fly into any appropriate coordinate system. As a result, it's possible to integrate and combine information from multiple GIS layers. This is a fundamental GIS capability. Accurate location forms the basis for almost all GIS operations.

About map projections

Whether you treat the earth as a sphere or a spheroid, you must transform its three-dimensional surface to create a flat map sheet. This mathematical transformation is commonly referred to as a map projection. One easy way to understand how map projections alter spatial properties is to visualize shining a light through the earth onto a surface, called the projection surface. Imagine the earth's surface is clear with the graticule drawn on it. Wrap a piece of paper around the earth. A light at the center of the earth will cast the shadows of the graticule onto the piece of paper. You can now unwrap the paper and lay it flat. The shape of the graticule on the flat paper is different from that on the earth. The map projection has distorted the graticule.

A spheroid can't be flattened to a plane any more easily than a piece of orange peel can be flattened—it will rip. Representing the earth's surface in two dimensions causes distortion in the shape, area, distance, or direction of the data.

A map projection uses mathematical formulas to relate spherical coordinates on the globe to flat, planar coordinates. Different projections cause different types of distortions. Some projections are designed to minimize the distortion of one or two of the data's characteristics. A projection could maintain the area of a feature but alter its shape. In the graphic below, data near the poles is stretched.

The following diagram shows how three-dimensional features are compressed to fit onto a flat surface.

Map projections are designed for specific purposes. One map projection might be used for large-scale data in a limited area, while another is used for a small-scale map of the world. Map projections designed for small-scale data are usually based on spherical rather than spheroidal geographic coordinate systems.

Conformal projections

Conformal projections preserve local shape. To preserve individual angles describing the spatial relationships, a Conformal projection must show the perpendicular graticule lines intersecting at 90-degree angles on the map. A map projection accomplishes this by maintaining all angles. The drawback is that the area enclosed by a series of arcs may be greatly distorted in the process. No map projection can preserve shapes of larger regions.

Equal area projections

Equal area projections preserve the area of displayed features. To do this, the other properties—shape, angle, and scale—are distorted. In Equal area projections, the meridians and parallels may not intersect at right angles. In some instances, especially maps of smaller regions, shapes are not obviously distorted, and distinguishing an Equal area projection from a Conformal projection is difficult unless documented or measured.

Equidistant projections

Equidistant maps preserve the distances between certain points. Scale is not maintained correctly by any projection throughout an entire map. However, there are in most cases, one or more lines on a map along which scale is maintained correctly. Most Equidistant projections have one or more lines in which the length of the line on a map is the same length (at map scale) as the same line on the globe, regardless of whether it is a great or small circle, or straight or curved. Such distances are said to be true. For example, in the Sinusoidal projection, the equator and all parallels are their true lengths. In other Equidistant projections, the equator and all meridians are true. Still others (for example, Two-point Equidistant) show true scale between one or two points and every other point on the map. Keep in mind that no projection is equidistant to and from all points on a map.

True-direction projections

The shortest route between two points on a curved surface such as the earth is along the spherical equivalent of a straight line on a flat surface. That is the great circle on which the two points lie. True-direction, or Azimuthal projections maintain some of the great circle arcs, giving the directions or azimuths of all points on the map correctly with respect to the center. Some True-direction projections are also conformal, equal area, or equidistant.

Projection types: illustrated

Each of the main projection types—Conic, Cylindrical, and Planar are illustrated below.

Conic (tangent)

Conic (secant)

CylindricalAspects

Planar Aspects

Polar Aspect (different perspectives)

SPATIALDATA

Also known as *geospatial data* or *geographic information* it is the data or information that identifies the geographic location of features and boundaries on Earth, such as natural or constructed features, oceans, and more. Spatial data is usually stored as coordinates and topology, and is data that can be mapped. Spatial data is often accessed, manipulated or analyzed through Geographic Information Systems (GIS).

ATRIBUTE DATA

Attribute data are descriptions, measurements, and/or classifications of geographic features in a map. Attribute data can be classified into 4 levels of measurement: nominal, ordinal, interval and ratio.

Database management system (**DBMS**) is a computer software application that interacts with the user, other applications, and the database itself to capture and analyze data. A general-purpose DBMS is designed to allow the definition, creation, querying, update, and administration of databases.

FUNCTIONS OF DBMS

A DBMS performs several important functions that guarantee integrity and consistency of data in the database. Most of these functions are transparent to end-users. There are the following important functions and services provided by a DBMS:

(i) Data Storage Management: It provides a mechanism for management of permanent storage of the data. The internal schema defines how the data should be stored by the storage management mechanism and the storage manager interfaces with the operating system to access the physical storage.

(ii) Data Manipulation Management: A DBMS furnishes users with the ability to retrieve, update and delete existing data in the database.

(iii) Data Definition Services: The DBMS accepts the data definitions such as external schema, the conceptual schema, the internal schema, and all the associated mappings in source form.

(iv) Data Dictionary/System Catalog Management: The DBMS provides a data dictionary or system catalog function in which descriptions of data items are stored and which is accessible to users.

(v) Database Communication Interfaces: The end-user's requests for database access are transmitted to DBMS in the form of communication messages.

(vi) Authorization / Security Management: The DBMS protects the database against unauthorized access, either international or accidental. It furnishes mechanism to ensure that only authorized users an access the database.

(vii) Backup and Recovery Management: The DBMS provides mechanisms for backing up data periodically and recovering from different types of failures. This prevents the loss of data,

(viii) Concurrency Control Service: Since DBMSs support sharing of data among multiple users, they must provide a mechanism for managing concurrent access to the database. DBMSs ensure that the database kept in consistent state and that integrity of the data is preserved.

(ix) Transaction Management: A transaction is a series of database operations, carried out by a single user or application program, which accesses or changes the contents of the database. Therefore, a DBMS must provide a mechanism to ensure either that all the updates corresponding to a given transaction are made or that none of them is made.

(x) Database Access and Application Programming Interfaces: All DBMS provide interface to enable applications to use DBMS services. They provide data access via Structured Query Language (SQL). The DBMS query language contains two components: (a) a Data Definition Language (DDL) and (b) a Data Manipulation Language (DML).

UNIT-5

Data Entry, Storage and Analysis

Syllabus:

Data models – vector and raster data – data compression – data input by digitization and scanning – attribute data analysis – integrated data analysis – Modeling in GIS Highway alignment studies, Highway Alignment, Factor affecting Highway Alignment – Land Information System.-importance.

Learning Outcomes:

- Understand the different data inputs and data types.
- Discuss the different data models and data analysis.

Spatial Data Models

Spatial data structures provide the information that the computer requires to reconstruct the spatial data model in digital form. Although some lines act alone and contain specific attribute information that describes their character, other more complex collections of lines called networks add a dimension of attribute characters. Thus not only does a road network contain information about the type of road or similar variables, but it will also indicate, that travel is possible only in a particular direction. This information must be extended to each connecting line segment to advise the user that movement can continue along each segment until the attributes change-perhaps until a one-way street becomes a two-way street. For example, one node might indicate the existence of a stop sign, a traffic signal, or a sign prohibiting U-turns. All these attributes must be connected throughout the network so that the computer knows the inherent real-world relationships that are being modelled within the network. Such explicit information about connectivity and relative spatial relationships is called topology.

vector data model: A representation of the world using points, lines, and polygons. Vector models are useful for storing data that has discrete boundaries, such as country borders, land parcels, and streets.

raster data model: A representation of the world as a surface divided into a regular grid of cells. Raster models are useful for storing data that varies continuously, as in an aerial photograph, a satellite image, a surface of chemical concentrations, or an elevation surface

Types of Raster GIS Models

The grid based GIS spatial data can be stored, manipulated, analysed, and referenced basically in anyone of the three methods/models. These three models (Burrough, 1983) are: GRID/LUNAR/MAGI model, IMGRID model and MAP model. All of these models use the grid cell values, their attributes, coverages and corresponding legends. These models are developed depending upon the requirements from time to time. Based on the applications of interest, availability of software's and other related information, anyone of the above models can be selected for the execution of a particular GIS project. There are a number of ways of forcing a computer to store and reference the individual grid cell values, their attributes, coverage names and legends.

GRID Model

The first and foremost model for the representation of raster data is the GIRD model. The method of storing, manipulating, and analysing the grid based data was first conceptualised by an attempt to deve'lop GRID model. Burrough (1983) used this approach, because each of those early GIS systems used this model. Fig. 8.9 (a) illustrates the GRID model. In this method, each grid cell is referenced and addressed individually and is associated with identically positioned grid cells in all other coverages, rather like a vertical column of grid cells, each dealing with a separate theme. Comparisons between coverages are therefore performed on a single column at a time. For example, to compare soil attributes in one coverage with vegetation attributes in a second coverage, land use/land cover attributes in a third coverage, each X and Y location must be examined individually. So a soil grid cell at location must be examined individually. So a soil grid cell at location X10-Y10 will be compared to its vegetation counterpart and third layer land use/land cover at location X 10-Y1 O. You might be able to envision this by imagining a geological core in which each rock type is lying directly on top of the next, and to get a picture of the entire study area, it will be necessary to put a large number of cores together.

The advantage of this model is that computational comparison of multiple themes or coverages for each grid cell location is relatively easy. This is a reasonable approach and has proven successful. The main disadvantage is that it limits the efficient examination of relationships of themes to one-to-one relationships within the spatial framework. In other words, it is more inconvenient to compare groups in one coverage to groups in another coverage because each grid cell location must be addressed individually. Second disadvantage is more storage space for the cell data and the representation is vertical rather than horizontal, which would more closely resemble our notion of maps.

IMGRID Model

With a slight modification of the checkerboard analog, the second basic raster data model, that is the IMGRID data model, can be illustrated (Fig. 8.9 (b». This model is also used in the early GIS system (Burrough, 1983). Let us assume that the red squares on checkerboard map serve to contain a single attribute, rather than just a theme. Instead, we can use the number 1 (red squares) to repre·sent water and 0 (black squares) to indicate the absence of water. How can we represent a thematic map of land use that contains, say four categories, namely, recreation, agriculture, industry, and residences? Each of these four attributes would have to be separated out as an individual layer. One layer would stand for agriculture only, with 1 's and O's representing the presence or absence of this activity for each grid cell. Recreation, industry, and residences would be represented in the same way, with each variable referenced directly, rather than referencing the grid cell as we did in the GRID/LUNAR/ MAGI data model. Finally, the coverages would be combined vertically, or in column fashion, to produce a single theme or coverage, much as red, yellow, green, and blue printing plates are combined to create a single colour image.

IMGRID system has two major advantages. First, we have a contiguous object that more closely resembles how we think about a map. That is, our primary storage object is a two-dimensional array of numbers, rather than a column of numbers for different themes. Second, we reduce the numbers that must be contained in each coverage to O's and 1 'so This will certainly simplify our computations and will eliminate the need for map legends. Since each variable is uniquely identified, assigning a single attribute value to a single grid cell is possible, and this is a third advantage. Let us assume that a given grid cell partly occupies agriculture and partly recreation and each of these attributes of land use theme is separated out. In such a case, we may encounter difficulties when creating our final thematic coverage if multiple values occur in individual cells. To avoid such problems, we must be able to ensure that each grid cell has only a single value for each variable.

The IMGRID model is seems to be more intuitive from a map abstraction viewpoint, and requires us to be very specific about the attributes to be contained in each coverage. But it offers the advantage of using the coverage as the direct object of reference for the computer. Its limitations stem primarily from the problem of data explosion. Imagine for a moment that you have a database composed of 50 themes. Each theme must be separated out into binary (O's and 1 's) coverages on the basis of individual attributes within each theme. Suppose that there is an average of 10 categories for each theme. To represent this rather modest database, you will need a total of 10 x 50 or 500 coverages. Although available storage devices can certainly manage such volumes, you need to manage and keep track of examining this approach further. Imagine how many values must be modified and recoded to create a new theme. For example, to combine 10 binary coverages to create a new thematic coverage with 10 categories, you would have to separate the thematic coverage into 20 new binary coverages each. Thus, for a simple operation you had to combine 10 grid cell values, and to create additional thematic coverage it is necessary to produce 10 new values of 0 and 1 for each variable. This is a rather tedious approach.

MAP Model

The third raster GIS model Map Analysis Package (MAP) model developed by C. Dana Tomlin (Burrough, 1983) formally integrates the advantages of the above two raster data structure methods. In this data model each thematic coverage is recorded and accessed separately by map name or title. This is accomplished by recording each variable, or mapping unit, of the coverage's theme as a separate number code or label, which can be accessed individually when the coverage is retrieved. The label corresponds to a portion of the legend and has its own symbol assigned to it. In this way, it is easy to perform operation on individual grid cells and groups of similar grid cells, and the resolution changes in value require rewriting only a single number per mapping unit, thus simplifying the computations. The overall major improvement is that the MAP method allows ready manipulation of the data in a many-to-one relationship of the attribute values and the sets of grid cells.

The MAP data model is compatible to almost all computer systems from its original mainframe version to Macintosh and PC versions and modern UNIX-based workstation versions. It can be used as a teaching version of GIS as it is very flexible and also becomes a major module in commercial GIS packages like ARC/INFO.

Although raster GIS systems have traditionally been developed to allow single attributes to be stored individually for each grid cell, some have evolved to include direct links to existing database management systems. This approach extends the utility of the raster GIS by minimising the number of coverages and substituting multiple variables for each grid cell in each coverage. Such extensions to the raster data model have also allowed direct linkage to existing GIS systems that use a vector back and forth from raster to vector. The user can operate with all the advantages of both the data structures. The conversion process is often quite transparent, allowing the user to perform the analyses needed without concern for the original data structure. This feature is particularly important because it is strengthening the relationship between traditional digital image processing software used to manipulate grid cellbased, remotely sensed data and GIS software. Many software systems already have both sets of capabilities, and still more are likely in the future. Together with the linkage with existing statistical packages, we are rapidly approaching the systems that operate with a superset of spatial analytical techniques, resulting in a maturing of automated geography.

Vector GIS Models

Vector data structures allow the representation of geographic space in an intuitive way reminiscent of the familiar analog map. The geographic space can be represented by the spatial location of items or attributes which are stored in another file for later access. Fig. 8.5 shows how the different entity, namely, points, lines, and areas can be defined by coordinate geometry. Like the raster spatial data model, there are many potential vector data models that can be used to store the geometric representation of entities in the computer.

The two basic types of vector data models are (i) spaghetti model, and (ii) topological model.

Spaghetti Model

The simplest vector data structure that can be used to reproduce a geographical image in the computer is a file containing (x, y) coordinate pairs that represent the location of individual point features. The shortest spaghetti can be represented as a point, collection of a number of point spaghettis for a line entity and collections of line segments that come together at the beginning and ending of surrounding areas form an area entity. Each entity is a single, logical record in the computer, coded as variable length strings of (x, y) coordinate pairs. Let us assume that two polygons lie adjacent to each other in a thematic coverage. These two adjacent polygons must have separate pieces of spaghetti for adjacent sides. That is, no two adjacent polygons share the same string of spaghetti. Each side of polygon is uniquely defined by its own set of lines and coordinate pairs. In this model of representing vector data, all the spaghetties are recorded separately for polygons. But in the computer they should have the same coordinates.

Topological Models

In order to use the data manipulation and analysis subsystem more efficiently and obtain the desired results, to allow advanced analytical techniques on GIS data and its systematic study in any project area, much explicit spatial information is to be created. The topological data model incorporates solutions to some of the frequently used operations in advanced GIS analytical techniques. This is done by explicitly recording adjacency information into the basic logical entity in topological data structures, beginning and ending when it contacts or intersects another line, or when there is a change in the direction of the line. Each line then has two sets of numbers: a pair of coordinates. and an associated node number. The node is the intersection of two or more lines, and its number is used to refer to any line to which it is connected. In addition, each line segment, called a link, has its own identification number that is used as a pointer to indicate the set of nodes that represent its beginning and ending polygon. These links also have identification codes that relate polygon numbers to see which two polygons are adjacent to each other along its length. In fact, the left and right polygon are also stored explicitly, so that even this tedious step is eliminated.

There are a number of topological vector data models. Out of the available models, three models are very common in use. These three models are: (a) GBFI DIME model created by US Department of Commerce, Bureau of the Census, 1969 (b) TIGER model (Marx, 1986) and (c) POLYVERT(Peuquet, 1984).

GBF/DIME Topological Vector Model

The best-known topological data model is the GBF/DIME (Geographical Base File/Duallndependent Map Encoding) model created by the US Bureau of the Census to automate the storage of street map data for the decennial census (US Department of Commerce, Bureau of the Census, 1969). GBF/DIME models were designed to incorporate topological information about urban areas for use in demographic analyses (Cooke, 1987) and were created by graph theory. In this case the straight-line segment ends when it either changes direction or intersects another line, and the nodes are identified with codes. In addition to the basic topological model, the GBF/DIME model assigns a directional code in'the form of a 'From node and a To node,' that is, a low-value node to a high-value node in the sequence.

TIGER Topological Vector Model

TIGER stands for Topologically Integrated Geographic Encoding and Referencing system. This model does not depend upon the graph theory designed for use in the 1990 US census. In this system, points, lines, and areas can be explicitly addressed, and therefore census blocks can be retrieved directly by block number rather than by relying on the adjacency information contained in the links. Real-world features such as meandering streams and irregular coastlines are given a graphic portrayal more representative of their true geographic shape. Thus TIGER files are more generally used in research which is not related to census.

POL YVRT Topological Vector Model

POLYVRT developed by Peucker and Chrisman (1975) and later implemented at the Harvard

Laboratory for Computer Graphics was called the POL YVRT (POLYgon con VERT) model. In this method of representing vector data, each type of geographic entity is stored separately. These separate objects are then linked in a hierarchical data structure with pOints relating to lines, which in turn are related to polygons through the use of pOinters. Each collection of line segments, is collectively called chains in this explicit directional information in the form of To-From nodes as well as left-right polygons.

RASTER AND VECTOR DATA

Raster data is made up of pixels (or cells), and each pixel has an associated value. Simplifying slightly, a digital photograph is an example of a raster dataset where each pixel value corresponds to a particular colour. In GIS, the pixel values may represent elevation above sea level, or chemical concentrations, or rainfall etc. The key point is that all of this data is represented as a grid of (usually square) cells. The difference between a digital elevation model (DEM) in GIS and a digital photograph is that the DEM includes additional information describing *where* the edges of the image are located in the real world, together with how big each cell is on the ground. This means that your GIS can position your raster images (DEM, hillshade, slope map etc.) correctly relative to one another, and this allows you to build up your map.

Vector data consists of individual points, which (for 2D data) are stored as pairs of (x, y) co-ordinates. The points may be joined in a particular order to create lines, or joined into closed rings to create polygons, but all vector data fundamentally consists of lists of co-ordinates that define vertices, together with rules to determine whether and how those vertices are joined.

		Note that whereas raster data consists of an array of regularly spaced cells, the points in a vector dataset need not be regularly spaced. In many cases, both vector and raster representations of the same data are possible: At this scale, there is very little difference between the vector representation and the "fine" (small pixel size) raster representation. However, if you zoomed in closely, you'd see the polygon edges of the fine raster would start to become pixelated, whereas the vector representation would remain crisp. In the "coarse" raster the pixelation is already clearly visible, even at this scale.

Vector Data

Advantages : Data can be represented at its original resolution and form without generalization. Graphic output is usually more aesthetically pleasing (traditional cartographic representation); Since most data, e.g. hard copy maps, is in vector form no data conversion is required. Accurate geographic location of data is maintained. Allows for efficient encoding of topology, and as a result more efficient operations that require topological information, e.g. proximity, network analysis.

Disadvantages: The location of each vertex needs to be stored explicitly. For effective analysis, vector data must be converted into a topological structure. This is often processing intensive and usually requires extensive data cleaning. As well, topology is static, and any updating or editing of the vector data requires re-building of the topology. Algorithms for manipulative and analysis functions are complex and may be processing intensive. Often, this inherently limits the functionality for large data sets, e.g. a large number of features. Continuous data, such as elevation data, is not effectively represented in vector form. Usually substantial data generalization or interpolation is required for these data layers. Spatial analysis and filtering within polygons is impossible.

Raster Data

Advantages : The geographic location of each cell is implied by its position in the cell matrix. Accordingly, other than an origin point, e.g. bottom left corner, no geographic coordinates are stored. Due to the nature of the data storage technique data analysis is usually easy to program and quick to perform. The inherent nature of raster maps, e.g. one attribute maps, is ideally suited for mathematical modeling and quantitative analysis. Discrete data, e.g. forestry stands, is accommodated equally well as continuous data, e.g. elevation data, and facilitates the integrating of the two data types. Grid-cell systems are very compatible with raster-based output devices, e.g. electrostatic plotters, graphic terminals.

Disadvantages: The cell size determines the resolution at which the data is represented.; It is especially difficult to adequately represent linear features depending on the cell resolution. Accordingly, network linkages are difficult to establish. Processing of associated attribute data may be cumbersome if large amounts of data exists. Raster maps inherently reflect only one attribute or characteristic for an area. Since most input data is in vector form, data must undergo vector-to-raster conversion. Besides increased processing requirements this may introduce data integrity concerns due to generalization and choice of inappropriate cell size. Most output maps from grid-cell systems do not conform to high-quality cartographic needs.

Vector Data types

Vector data is not made up of a grid of pixels. Instead, vector graphics are comprised of **vertices and paths**.

The three basic symbol types for vector data are points, lines and polygons (areas). Since the dawn of time, maps have been using symbols to represent real-world features. In GIS terminology, real-world features are called **spatial entities**.

The cartographer decides how much data needs to be generalized in a map. This depends on scale and how much detail will be displayed in the map. The decision to choose vector points, lines or polygons is governed by the cartographer and scale of the map.

Points

Point Vector Data Type: Simple XY Coordinates

Vector points are simply XY coordinates. When features are too small to be represented as polygons, points are used.

For example:

At a regional scale, city extents can be displayed as polygons because this amount of detail can be seen when zoomed in. But at a global scale, cities can be represented as points because the detail of city boundaries cannot be seen.

Vector data are stored as pairs of XY coordinates (latitude and longitude) represented as a point. Complementary information like street name or date of construction could accompany it in a table for its current use.

Lines

Vector Data Type Line: Connect the dots and it becomes a line feature

Vector lines connect vertices with paths. If you were to connect the dots in a particular order, you would end up with a **vector line** feature.

Lines usually represent features that are linear in nature. Cartographers can use a different thickness of line to show size of the feature. For example, 500 meter wide river may be thicker than a 50 meter wide river.

They can exist in the real-world such as roads or rivers. Or they can also be artificial divisions such as regional borders or administrative boundaries.

Points are simply pairs of XY coordinates (latitude and longitude). When you connect each point or vertex with a line in a particular order, they become a **vector line** feature.

Networks are line data sets but they are often considered to be different. This is because linear networks are topologically connected elements. They consist of junctions and turns with connectivity. If you were to find an optimal route using a traffic line network, it would follow one-way streets and turn restrictions to solve an analysis. Networks are just that smart.

Polygons

Vector Data Type Polygon: Connect the dots and enclose. It becomes a polygon feature

When a set of vertices are joined in a particular order and closed, they becomes a **vector polygon** feature. In order to create a polygon, the first and last coordinate pair are the same and all other pairs must be unique.

Polygons represent features that have a two-dimensional area. Examples of polygons are buildings, agricultural fields and discrete administrative areas.

Cartographers use polygons when the map scale is large enough to be represented as polygons.

Raster Spatial Data Types

Raster data is made up of pixels (also referred to as grid cells). They are usually regularly-spaced and square but they don't have to be. Rasters often look pixelated because each pixel is associated with a value or class.

For example:

Each pixel value in a digital photograph is associated with a red, green and blue value. Or each value in a digital elevation model represents a value of elevation. It could represent anything from thematic categories, heights or spectral value.

Raster models are useful for storing data that varies continuously, as in an aerial photograph, an elevation surface or a satellite image. But it depends on the cell size for spatial accuracy.

Raster data models can be discrete and continuous.

Discrete rasters

Discrete rasters have definable boundaries like this land cover classification raster.

Discrete rasters are also referred to as thematic or categorical raster data. They have distinct themes or categories. For example, one grid cell represents a land cover class or a soil type.

In a discrete raster land cover/use map, you can distinguish each thematic class. Each class can be discretely defined where it begins and ends. Each land cover cell is definable. The land cover class fills the entire area of the cell

Discrete data usually consists of integers to represent classes. For example, the value 1 might represent urban areas, the value 2 represents forest, etc. Political boundaries or ownership are other examples of discrete rasters.

Continuous Rasters

Continuous raster

Continuous rasters are grid cells with gradual changing data such as elevation, temperature or an aerial photograph. Continuous data is also known as non-discrete or surface data.

A continuous raster surface can be derived from a **fixed registration point**. For example, a digital elevation model is measured from sea level. Each cell represents a value above or below sea level. An aspect cell value is derived from a fixed direction such as north, east, south or west.

Phenomena can gradually vary along a continuous raster from a **specific source**. For example, a raster depicting an oil spill can show how the fluid moves from high concentration to low concentration. At the source of the oil spill, concentration is higher. It diffuses outwards with diminishing values as a function of distance.

Raster Data Compression

We can distinguish different ways of storing raster data, which basically vary in storage size and consequently in the geometric organisation of the storage. The following types of geometric elements are identified:

- Lines
- Stripes
- Tiles
- Areas (e.g. Quad-trees)
- Hierarchy of resolution

Raster data are managed easily in computers; all commonly used programming languages well support array handling. However, a raster when stored in a raw state with no compression can be extremely inefficient in terms of computer storage space.

As already said the way of improving raster space efficiency is data compression.

Illustrations and short texts are used to describe different methods of raster data storage and raster data compression techniques.

Full raster coding (no-compression)

By convention, raster data is normally stored row by row from the top left corner.

Runlength coding (lossless)

Geographical data tends to be "spatially autocorrelated", meaning that objects which are close to each other tend to have similar attributes:

"*All things are related, but nearby things are more related than distant things*" (Tobler 1970) Because of this principle, we expect neighboring pixels to have similar values. Therefore, instead of repeating pixel values, we can code the raster as pairs of numbers - (run length, value).

The runlength coding is a widely used compression technique for raster data. The primary data elements are pairs of values or tuples, consisting of a pixel value and a repetition count which specifies the number of pixels in the run. Data are built by reading successively row by row through the raster, creating a new tuple every time the pixel value changes or the end of the row is reached.

Describes the interior of an area by run-lengths, instead of the boundary.

In multiple attribute case there a more options available:

We can note in *Codes - III*, that if a run is not required to break at the end of each line we can compress data even further.

Chain coding (lossless)

See Rasterising vector data and the Freeman coding

Blockwise coding (lossless)

This method is a generalization of run-length encoding to two dimensions. Instead of sequences of 0s or 1s, square blocks are counted. For each square the position, the size and, the contents of the pixels are stored.

Quandtree coding (lossless)

The quadtree compression technique is the most common compression method applied to raster data. Quadtree coding stores the information by subdividing a square region into quadrants, each of which may be further subdivided in squares until the contents of the cells have the same values.

Example 2:

On the following figure we can see how an area is represented on a map and the corresponding quadtree representation. More information on constructing and addressing quadtrees, see Lesson "Spatial partitioning and indexing", Unit 2

Huffman coding (lossless compression)

Human coding compression technique involves preliminary analysis of the frequency of occurrence of symbols. Huffman technique creates, for each symbol, a binary data code, the length of which is inversely related to the frequency of occurrence.

DATAINPUT TECHNIQUES

Since the input of attribute data is usually quite simple, the discussion of data input techniques will be limited to spatial data only. There is no single method of entering the spatial data into a GIS. Rather, there are several, mutually compatible methods that can be used singly or in combination.

Digitizing

While considerable work has been done with newer technologies, the overwhelming majority of GIS spatial data entry is done by manual digitizing. A digitizer is an electronic device consisting of a table upon which the map or drawing is placed. The user traces the spatial features with a hand-held magnetic pen, often called a *mouse* or cursor. While tracing the features the coordinates of selected points, e.g. vertices, are sent to the computer and stored. All points that are recorded are registered against positional control points, usually the map corners, that are keyed in by the user at the beginning of the digitizing session. The coordinates are recorded in a user defined coordinate system or map projection. Latitude and longitude and UTM is most often used. The ability to adjust or transform data during digitizing from one projection to another is a desirable function of the GIS software. Numerous functional techniques exist to aid the operator in the digitizing process.

Digitizing can be done in a *point mode*, where single points are recorded one at a time, or in a *stream mode*, where a point is collected on regular intervals of time or distance, measured by an X and Y movement, e.g. every 3 metres. Digitizing can also be done blindly or with a graphics terminal. Blind digitizing infers that the graphic result is not immediately viewable to the person digitizing. Most systems display the digitized linework as it is being digitized on an accompanying graphics terminal.

Most GIS's use a *spaghetti mode* of digitizing. This allows the user to simply digitize lines by indicating a start point and an end point. Data can be captured in point or stream mode. However, some systems do allow the user to capture the data in an arc/node topological data structure. The arc/node data structure requires that the digitizer identify nodes.

Data capture in an arc/node approach helps to build a topologic data structure immediately. This lessens the amount of post processing required to *clean* and build the topological definitions. However, most often digitizing with an arc/node approach does not negate the requirement for editing and cleaning of the digitized linework before a complete topological structure can be obtained.

The building of topology is primarily a post-digitizing process that is commonly executed in *batch mode* after data has been cleaned. To date, only a few commercial vector GIS software offerings have successfully exhibited the capability to build topology interactively while the user digitizes.

Manual digitizing has many advantages. These include:

	Low capital cost, e.g. digitizing tables are cheap;
	Low cost of labour;
	Flexibility and adaptability to different data types and sources;
	Easily taught in a short amount of time - an easily mastered skill
	Generally the quality of data is high;
	Digitizing devices are very reliable and most often offer a greater precision that the data warrants; and
	Ability to easily register and update existing data.

For raster based GIS software data is still commonly digitized in a vector format and converted to a raster structure after the building of a clean topological structure. The procedure usually differs minimally from vector based software digitizing, other than some raster systems allow the user to define the resolution size of the grid-cell. Conversion to the raster structure may occur *on-the-fly* or afterwards as a separate conversion process.

Automatic Scanning

A variety of scanning devices exist for the automatic capture of spatial data. While several different technical approaches exist in scanning technology, all have the advantage of being able to capture spatial features from a map at a rapid rate of speed. However, as of yet, scanning has not proven to be a viable alternative for most GIS implementation. Scanners are generally expensive to acquire and operate. As well, most scanning devices have limitations with respect to the capture of selected features, e.g. text and symbol recognition. Experience has shown that most scanned data requires a substantial amount of manual editing to create a clean data layer. Given these basic constraints some other practical limitations of scanners should be identified. These include :

	hard copy maps are often unable to be removed to where a scanning device is available, e.g. most companies or agencies cannot afford their own scanning device and therefore must send their maps to a private firm for scanning;
	hard copy data may not be in a form that is viable for effective scanning, e.g. maps are of poor quality, or are in poor condition;
	geographic features may be too few on a single map to make it practical, cost-justifiable, to scan;
	often on *busy* maps a scanner may be unable to distinguish the features to be captured from the surrounding graphic information, e.g. dense contours with labels;
	with raster scanning there it is difficult to read unique labels (text) for a geographic feature effectively; and
	Scanning is much more expensive than manual digitizing, considering all the cost/performance issues.

Highway Alignment

Land Information System:

UNIT-6

RS & GIS APPLICATIONS

RS and GIS Applications:

Land cover and Land use, Agriculture, forestry, geology, geomorphology, urban applications, hydrology-flood zone delineation, and ground water prospectus, recharge reservoir storage estimation.

Learning Outcomes:

- To evaluate quick solutions and apply the same in solving problems using RS&GIS Techniques.

Land use land cover

Mapping of LULC and change detection using remote sensing and GIS techniques is a cost effective method of obtaining a clear understanding of the land cover alteration processes due to land use change and their consequences. Forest clearing has been identified as one of the most significant causes of deforestation in different parts of the world. Detailed scientific studies illustrate the apparent effect of farming activities resulting in modification of the original vegetation. The rate of deforestation is alarming in West Africa due to rapid population growth and land use.

To understand how LULC change affects and interact with global earth systems, information is needed on what changes occur, where and when they occur, the rate at which they occur, and the social and physical forces that drive those changes Moreover, there appears to be a gap in the available information and national decision-making process and rational planning. Therefore this paper seek to investigate (i.e. map out), identify and quantify the changes in LULC over the years.

Agricultural Applications:

GIS can be used to create more effective and efficient farming techniques. It can also analyze soil data and to determine: what are the best crop to plant?, where they should go? how to maintain nutrition levels to best benefit crop to plant?. It is fully integrated and widely accepted for helping government agencies to manage programs that support farmers and protect the environment. This could increase food production in different parts of the world so the world food crisis could be avoided.

GIS Applications in Geology:

Geologists use GIS in a various applications. The GIS is used to study geologic features, analyze soils and strata, assess seismic information, and or create three dimensional (3D) displays of geographic features. GIS can be also used to analyze rock information characteristics and identifying the best dam site location.

Irrigation water management: Water availability for irrigation purposes for any area is vital for crop production in that region. It needs to be properly and efficiently managed for the proper utilization of water. To evaluate the irrigation performance, integrated use of satellite remote sensing and GIS assisted by ground information has been found to be efficient technique in spatial and time domain for identification of major crops and their conditions, and determination of their areal extent and yield. Irrigation requirements of crop were determined by considering the factors such as evapotranspiration, Net Irrigation Requirement, Field irrigation Requirement, Gross Irrigation Requirement, and month total volume of water required, by organizing them in GIS environment.

GIS & RS Applications in Forestry:

GIS and RS technology is important in forest cover mapping.Multilayered site representation can be possible by using elevation (ex. slope), hydrology and the location information of roads and infrastructure this interpreted data of large areas can useful for Emergency and Fire mapping. And forest fires are directly affected on plants, animals, vegetation cover, stream flow, soil, air quality and climate. The loss the life and property also caused by loss of timber. Forest fires also responsible for wildlife habitat destruction. GIS useful for forest fire modelling obtaining timely and efficient geospatial information for management of forest fires .

GIS data useful for forest management, because most of the rainforest are depleting in enormous rate, and it is due to the increasing rate of urbanization and agriculture and this human activities encroachment in forest areas.

Deforestation Identification using different RS data i.e. different multispectral imageries with GIS maps of the urban growth and factor recognition affecting on deforestation. Also Different temporal satellite image data can provides statistical analysis of forest cover and information about deforestation using GIS maps.
GIS is useful for representation in the form of graphs, maps and other GIS statistical modelling functionalities aids its value. So it is useful for forest management
DEM (Digital Elevation Data) of forest cover useful for GIS analysis. And it is useful for various terrain attributes examination, movement of soil and nutrients influence from it, as well as the resulting outcome on wildlife productivity, forest, plant distribution.