Chapter 5

Principles of Satellite Remote Sensing

Goal: Give a overview on the characteristics of satellite remote sensing.

Satellites have several unique characteristics which make them particularly useful for remote sensing of the Earth's surface and atmosphere.

5.1 Satellite Characteristics: Orbits

The path followed by a satellite is a circular orbit. Satellite orbits are matched to the capability and objective of the sensor(s) they carry. Orbit selection can vary in terms of altitude (their height above the Earth's surface) and their orientation and rotation relative to the Earth. The circular orbits are predicted by matching the centripetal acceleration to the gravitational force.

For a circular satellite orbit around a spherically homogenous planet the gravitational force F_g and centrifugal force F_c are in balance:

$$F_g = mg\left(\frac{R}{r}\right)^2 \tag{5.1}$$

$$F_c = \frac{mv^2}{r} = m\omega^2 r \tag{5.2}$$

with: mass of satellite m, orbit height h, orbit radius r, speed v, angular velocity $\omega = v/r$ and for the earth with gravitational constant $g = 9.81\,\mathrm{m/s^2}$

and $R = 6380\,\mathrm{km}$, this gives us for a circular orbit:

$$F_g = F_c \Rightarrow g\left(\frac{R}{r}\right)^2 = \frac{v^2}{r} \tag{5.3}$$

$$v = \sqrt{\frac{gR^2}{r}} \tag{5.4}$$

Note: no dependence on satellite mass!
According to that the orbital period is given by:

$$T = \frac{2\pi r}{v} = 2\pi r\sqrt{\frac{r}{gR^2}} \tag{5.5}$$

Equation 4.5 predicts the higher the orbit, the longer its period of rotation

---

Examples:

Example 1: geostationary orbit (e.g. communication satellites) T = 23h56 min; R = 6380km $r = \left(\frac{TR\sqrt{g}}{2\pi}\right)^{2/3}$ $h = r - R = ... =$ 35808km (height of satellite orbit)

Example 2: many remote sensing satellites: h = 800km v = 7.46km/s, T $\Rightarrow$ ca.101 min.

Note: T dictates h and vice versa

Satellites used to observe the Earth's atmosphere and surface fly principally in one of the two orbits, geostationary Earth orbit (GEO) or low Earth orbit (LEO).

5.2 Geostationary Orbits (GEO)

If a satellite is moved to an orbit enough out from Earth then the period of this orbit can be selected to match the rotation of the Earth. This synchronization means that for an observer at a fixed location on Earth, a satellite in a geosynchronous orbit returns to exactly the same place in the sky at exactly the same time each day. The special case of a geosynchronous orbit that is circular and directly above the equator is called a geostationary orbit (GEO). A GEO has an orbital eccentricity of zero. A satellite in that orbit, at that chosen distance from Earth in the equatorial plane, moving in a counter-clockwise sense, moves around the Earth exactly the same speed as the Earth rotates. The satellite then stays fixed in the same place above the Earth in a geostationary orbit (see above calculating its altitude). This allows the satellites to observe and collect information continuously over specific areas. From the ground, a geostationary object appears motionless in the sky.

Advantages of GEO:

- See whole Earth disk at once due to large distance
- See same spot on the surface all the time i.e. high temporal coverage

For weather monitoring satellites which require critical information on atmospheric dynamics for short-term forecasting and numerical weather prediction (NWP) these satellite have the best prerequisites. Weather and communications satellites commonly have these types of orbits. To get temporal resolution big advantage for weather monitoring satellites (e.g. Meteosat monitores the whole earth disk every 15 min) because knowing atmospheric dynamics is critical to short-term forecasting and numerical weather prediction (NWP). With the combination of several GEO meteorological satellites the weather of the whole Earth can be monitored!

Satellites in geostationary orbit must all occupy a single ring above the equator. The requirement to space these satellites apart means that there are a limited number of orbital "slots" available, thus only a limited number of satellites can be placed in geostationary orbit. This has led to conflict between different countries wishing access to the same orbital slots (countries at the same longitude but differing latitudes). Due to the constant 0° latitude, satellite locations may differ by longitude only.

Disadvantages of GEO:

- Typically low spatial resolution due to high altitude: e.g. METEOSAT 2nd Generation (MSG) 1 x 1 km visible, 3 x 3 km IR (used to be 3 x 3 & 6 x 6 respectively)
- Spatial resolution at 60-70° several times lower

- not much usuable information beyond 70°, cannot see poles very well (since the orbit is over the equator)

Side remark: A semisynchronous orbit has an orbital period of 0.5 sidereal days, i.e. 11h 58min. Relative to the Earth's surface it has twice this period. Examples include the Molniya orbit and the orbits of the satellites in the Global Positioning System. These are often used for GPS applications.

5.3 Lower Earth Orbits (LEO)

Lower Earth orbits (LEO) are between 300 and 2000 km altitude, typically below 1000 km and generally give higher spatial resolution and launch costs are cheaper than for GEO satellites.There are two types of LEO: (Near)-polar orbits and inclination orbits. The disadvantages of LEO satellites are that they need to launch to precise the altitude and orbital inclination and that the orbital decay at LEOs (Low Earth Orbits) ¡ 1000km is limiting their lifetime because from the drag from atmosphere causes the orbit to become more eccentric. The drag increases with increasing solar activity (sun spots).

5.3.1 Polar and Near Polar Orbits

Polar orbiting satellites overfly higher latitudes even though the satellite does not pass directly over the poles. The altitude of polar orbiting satellites is normally lower than 2000 km. Many remote sensing platforms are designed to follow an orbit (basically north-south) which, in conjunction with the Earth's rotation (west-east), allows them to cover most of the Earth's surface over a certain period of time. These are nearpolar orbits, so named for the inclination of the orbit relative to a line running between the North and South poles.

Characteristics of polar and near polar orbits:

- full polar orbit inclined 90° to equator
- typically few degrees off (larger than 90°), so poles not covered

- orbital period, T, typically 90 to 110 min
- near circular orbit between 300 km and 1000 km (low Earth orbit)
- typically higher spatial resolution than geostationary
- rotation of Earth under satellite gives (potential) total coverage

I =**Angle of inclination:**

- Determines the poleward extent of the orbit
- Defined as the angle measured in the counter-clockwise direction between the equatorial plane and the plane of the orbit

Nodes of the orbit:

- The point of the orbit where the satellite crosses the equator:
- Moving northerly direction: ascending
- Moving southerly direction: descending

Sunsynchronous orbits: Many of the LEO near-polar satellite orbits (700 to 1000 km) are also sun-synchronous such that they cover each area of the world at a constant local time of day called local sun time. At any given latitude, the position of the sun in the sky as the satellite passes overhead will be the same within the same season. This ensures consistent illumination conditions when acquiring images in a specific season over successive years, or over a particular area over a series of days. This is an important factor for monitoring changes between images or for mosaicking adjacent images together, as they do not have to be corrected for different illumination conditions.

Explanation: The mechanism for providing this precession of the orbit is achieved using the effects of the non-uniformity of the Earth gravitational field that arises from the fact that the Earth is not a prefect sphere.

- The Earth is not spherical (gravitation potential = $1/(r$+higher order terms) which leads to precession of the orbit if orbit not in equator plane $\Rightarrow$ the orbital node changes due to the precession.

The rate of change for the orbital node (precession rate - $d\Omega/dt$) is approximately given by:

$$\frac{d\Omega}{dt} = -\frac{3}{2} J_2 R^3 \sqrt{g} \frac{\cos I}{r^{7/2}} \tag{5.6}$$

Here $J_2 = 0.00108$ which is the second harmonic of the Earth geopotential and I ist the inclination angle.

- precession rate = function of I and r
- for $I = 98°$, precession period is 1 year! I.e., the position of the orbit relative to the sun is constant $\Rightarrow$ satellite passes over same latitude at the same local time $\Rightarrow$ useful to avoid varying illumination conditions caused by different time of day $\Rightarrow$ used by most remote-sensing satellites
- Note: seasonal variation of illumination cannot be avoided!
- ground track of sun-synchronous, near polar orbiting satellite: wavy line

Repeat cycle: After a number of orbits, sub-satellite point retraces its path (for ENVISAT 35 days)

Revisit period: The satellite revisit time is the time elapsed between observations of the same point on earth by a satellite: it depends on the satellite's orbit, target location, and swath of the sensor. In near-polar orbits, areas at high latitudes will be imaged more frequently than the equatorial zone due to the increasing overlap in adjacent swaths as the orbit paths come closer together near the poles.

If the orbit is also sunsynchronous, the ascending pass is most likely on the shadowed side of the Earth while the descending pass is on the sunlit side. Sensors recording reflected solar energy only image the surface on a descending pass, when solar illumination is available. Active sensors which provide their own illumination or passive sensors that record emitted (e.g. thermal) radiation can also image the surface on ascending.

5.3.2 Tropical or Inclination orbits

- Orbits with inclination angles greater than 0° and lower than 90°

- Determined by the region on Earth that is of most interest (i.e., an instrument to study the tropics may be best put on a low inclination satellite)
- Orbital altitude of these satellites generally on the order of a few hundred km
- Satellites are not sun-synchronous - view a place on Earth at varying times

Advantages of LEO:

- typically higher spatial resolution than geostationary
- rotation of Earth under satellite gives (potential) global coverage
- launch costs much cheaper than GEO

Disadvantages of LEO:

- need to launch to precise altitude and orbital inclination
- orbital decay at LEO ¡1000 km because drag from the atmosphere causes the orbit to become more eccentric. (Drag increases with increasing solar activity (sun spots)

5.4 Spatial resolution

As a satellite revolves around the Earth, the sensor "sees" a certain portion of the Earth's surface. The area imaged on the surface, is referred to as the swath. Imaging swaths for spaceborne sensors generally vary between tens and hundreds of kilometres wide. As the satellite orbits the Earth from pole to pole, its east-west position would not change if the Earth did not rotate.

However, as seen from the Earth, it seems that the satellite is shifting westward because the Earth is rotating (from west to east) beneath it. This apparent movement allows the satellite swath to cover a new area with each consecutive pass. The satellite's orbit and the rotation of the Earth work together to allow complete coverage of the Earth's surface, after it has completed one complete cycle of orbits.

The platform, plays a large role in determining the detail of information obtained and the total area imaged by the sensor. Sensors onboard platforms far away from their targets, typically view a larger area, but cannot provide great detail. Compare what an astronaut onboard the space shuttle sees of the Earth to what you can see from an airplane. The astronaut might see your whole province or country in one glance, but couldn't distinguish individual houses. Flying over a city or town, you would be able to see individual buildings and cars, but you would be viewing a much smaller area than the astronaut. There is a similar difference between satellite images and airphotos. The detail discernible in an image is dependent on the spatial resolution of the sensor and refers to the size of the smallest possible feature that can be detected.

Spatial resolution of passive sensors (we will look at the special case of active microwave sensors later) depends primarily on their Instantaneous Field of View ($IFOV$) and the height of the satellite orbit (h). The $IFOV$ is the angular cone of visibility of the sensor and determines the area on the Earth's surface which is "seen" from a given altitude at one particular moment in time. The size of the area viewed is determined by multiplying the IFOV by the distance from the ground to the sensor:

$$GIFOV = h * \tan(IFOV) \tag{5.7}$$

$IFOV$ depends on sensors optic and wavelength. This area on the ground is called the Ground-Projected $IFOV$ ($GIFOV$)resolution cell and determines a sensor's maximum spatial resolution. For a homogeneous feature to be detected, its size generally has to be equal to or larger than the $GIFOV$. If the feature is smaller than this, it may not be detectable as the average brightness of all features in that $GIFOV$ will be recorded. However, smaller features may sometimes be detectable if their reflectance dominates within a particular $GIFOV$ allowing sub-pixel or -$GIFOV$ detection.

Most remote sensing images are composed of a matrix of picture elements, or pixels, which are the smallest units of an image. Image pixels are normally square and represent a certain area on an image. It is important to distinguish between pixel size and spatial resolution - they are not interchangeable. If a sensor has a spatial resolution of 20 m and an image from that sensor is displayed at full resolution, each pixel represents an area of 20 m x 20 m on the ground. In this case the pixel size and resolution are the same. However, it is possible to display an image with a pixel size different than the resolution. Many posters of satellite images of the Earth have their pixels averaged to represent larger areas, although the original spatial resolution of the sensor that collected the imagery remains the same.

Pixels = smallest units of an image Full resolution: pixel size = spatial resolution

Images where only large features are visible are said to have coarse or low resolution. In fine or high resolution images, small objects can be detected. Military sensors for example, are designed to view as much detail as possible, and therefore have very fine resolution. Commercial satellites provide imagery with resolutions varying from a few metres to several kilometres. Generally speaking, the finer the resolution, the less total ground area can be seen.

The ratio of distance on an image or map, to actual ground distance is referred to as scale. Maps or images with small ”map-to-ground ratios” are referred to as small scale (e.g. 1:100,000), and those with larger ratios (e.g.1:5,000) are called large scale.

5.5 Temporal Resolution

In addition to spatial, spectral, and radiometric resolution, the concept of temporal resolution is also important to consider in a remote sensing system. The revisit period of a LEO satellite sensor is usually several days. Therefore the absolute temporal resolution of a remote sensing system to image the exact same area at the same viewing angle a second time is equal to this period. However, because of some degree of overlap in the imaging swaths of adjacent orbits for most satellites and the increase in this overlap with increasing latitude, some areas of the Earth tend to be re-imaged more frequently. Also, some satellite systems are able to point their sensors to image the same area between different satellite passes separated by periods from one to five days. Thus, the actual temporal resolution of a sensor depends on a variety of factors, including the satellite/sensor capabilities, the swath overlap, and latitude. The ability to collect imagery of the same area of the Earth’s surface at different periods of time is one of the most important elements for applying remote sensing data. Spectral characteristics of features may change over time and these changes can be detected by collecting and comparing multi-temporal imagery. For example, during the growing season, most species of vegetation are in a continual state of change and our ability to monitor those subtle changes using remote sensing is dependent on when and how frequently we collect imagery. By imaging on a continuing basis at different times we are able to monitor the changes that take place on the Earth's surface, whether they are naturally occurring (such as changes in natural vegetation cover or flooding) or induced by humans (such as urban development or deforestation). The time factor in imaging is important when:

- persistent clouds offer limited clear views of the Earth's surface (often in the tropics)
- short-lived phenomena (floods, oil slicks, etc.) need to be imaged
- multi-temporal comparisons are required (e.g. the spread of a forest disease from one year to the next)
- the changing appearance of a feature over time can be used to distinguish it from nearsimilar features (wheat / maize)

5.6 Spectral and Radiometric Resolution

Different classes of features and details in an image can often be distinguished by comparing their responses over distinct wavelength ranges. Broad classes, such as water and vegetation, can usually be separated using very broad wavelength ranges (within the visible and near infrared). Other more specific classes, such as different rock types, may not be easily distinguishable using either of these broad wavelength ranges and would require comparison at much finer wavelength ranges to separate them. Thus, a sensor is required with higher spectral resolution. Spectral resolution describes the ability of a sensor to define fine wavelength intervals. The finer the spectral resolution, the narrower the wavelength range for a particular channel or band.

Many remote sensing systems record energy over several separate wavelength ranges at various spectral resolutions. These are referred to as multi-spectral sensors and will be described in following lectures within the different applications. Advanced multi-spectral sensors called hyperspectral sensors, detect hundreds of very narrow spectral bands throughout the visible, near-infrared, and mid-infrared portions of the electromagnetic spectrum. Their very high spectral resolution facilitates fine discrimination between different targets based on their spectral response in each of the narrow bands.

While the arrangement of pixels describes the spatial structure of an image, the radiometric characteristics describe the actual information content in an image. Every time an image is acquired by a sensor, its sensitivity to the magnitude of the electromagnetic energy determines the radiometric resolution. The radiometric resolution of an imaging system describes its ability to discriminate very slight differences in energy. The finer the radiometric resolution of a sensor, the more sensitive it is to detecting small differences in reflected or emitted energy: Imagery data are represented by positive digital numbers which vary from 0 to (one less than) a selected power of 2. This range corresponds to the number of bits used for coding numbers in binary format. Each bit records an exponent of power 2 (e.g. 1 bit=2 1=2). The maximum number of brightness levels available depends on the number of bits used in representing the energy recorded. Thus, if a sensor used 8 bits to record the data, there would be $2^8 = 256$ digital values available, ranging from 0 to 255. However, if only 4 bits were used, then only $2^4 = 16$ values ranging from 0 to 15 would be available. Thus, the radiometric resolution would be much less. Image data are generally displayed in a range of grey tones, with black representing a digital number of 0 and white representing the maximum value (for example, 255 in 8-bit data). By comparing a 2-bit image with an 8-bit image, we can see that there is a large difference in the level of detail discernible depending on their radiometric resolutions.