Data Sources

Digital Image Data - Remote Sensing case:

- data of the earth's surface acquired from either aircraft or spacecraft platforms
- available in digital format;
- spatially the data is composed of discrete picture elements, or *pixels,*
- radiometrically it is quantised into discrete brightness levels.

Even data that are not recorded in digital form initially can be converted into discrete data by use of digitizing equipment such as scanning microdensitometers.

The great advantage of having data available digitally is that it can be processed by computer either for machine assisted information extraction or for embellishment before an image product is formed.

The computer is used to assist the role of photointerpretation.

A major characteristic of an image in remote sensing is the wavelength band it represents. Passive systems are working with images:

Some images are measurements of the **spatial disposition of reflected solar radiation** in the:
- ultraviolet,
- visible,
- near-to-middle infrared range of wavelengths.

Others are measurements of the **spatial distribution of energy emitted by the earth** itself (dominant in the so-called *thermal* infrared wavelength range).

Yet others, particularly in the **microwave band** of wavelengths, measure the relative return from the earth's surface of **energy actually transmitted from the vehicle** itself – these are active systems.

From a data handling and analysis point of view the properties of image data of **significance** are:

- the number and location of the spectral measurements (or spectral bands) provided by a particular sensor,

- the spatial resolution as described by the pixel size, in equivalent ground metres,
- the radiometric resolution, which describes the range and discernable number of discrete brightness values and is sometimes referred to alternatively as dynamic range or signal to noise ratio.

Frequently the radiometric resolution is expressed in terms of the number of binary digits, or bits, necessary to represent the range of available brightness values. Thus data with 8 bit radiometric resolution has 256 levels of brightness.
Together with the frame size of an image, in equivalent ground kilometres, the number of spectral bands, radiometric resolution and spatial resolution determine the data volume provided by a particular sensor and thus establish the amount of data to be processed, at least in principle.

As an illustration consider the **Landsat Thematic Mapper** instrument:

- it has 7 wavelength bands with 8 bit radiometric resolution,
- six of which have 30 m spatial resolution,
- and one of which has a spatial resolution of 120 m (the thermal band, for which the wavelength is so long that a larger aperture (anga, diapazonas) is necessary to collect sufficient signal energy to maintain the radiometric resolution).

An image frame of ***185 km x 185 km*** therefore represents ***2.37*** million pixels in the thermal band and ***38*** million pixels in the other six bands.

At 8 bits per pixel a complete 7 band image is composed of $\boldsymbol{1.848 \times 10^9}$ bits or ***1.848*** Gbit; alternatively and more commonly the data value would be expressed as ***231*** Mbytes.

Spectral Ranges Used in Remote Sensing

In principle, remote sensing systems could measure energy emanating from the carth's surface in any sensible range of wavelengths. However technological considerations, the selective ***opacity*** (nepermatomumas) of the earth's atmosphere, scattering from atmospheric particulates and the significance of the data provided exclude certain wavelengths.

The major ranges utilized for earth resources sensing are between about ***0.4 – 12 ìm*** (referred to as the *visible/infrared* range) and between about ***30 – 300 mm*** (referred to as the *microwave* range).

At microwave wavelengths it is often more **common to use** frequency rather than wavelength to describe ranges of importance. Thus the microwave range of 30 to 300 mm corresponds to frequencies between ***1 GHz – 10 GHz***.

For atmospheric remote sensing frequencies in the range ***20 GHz – 60 GHz*** are encountered.

The ***significance*** of these different ranges lies in the interaction mechanism between the electromagnetic radiation and the materials being interrogated.

In the ***visible/infrared*** range the reflected energy measured by a sensor depends upon properties:
- the pigmentation,
- moisture content
- cellular structure of vegetation,
- the mineral and moisture contents of soils,
- the level of sedimentation of water.

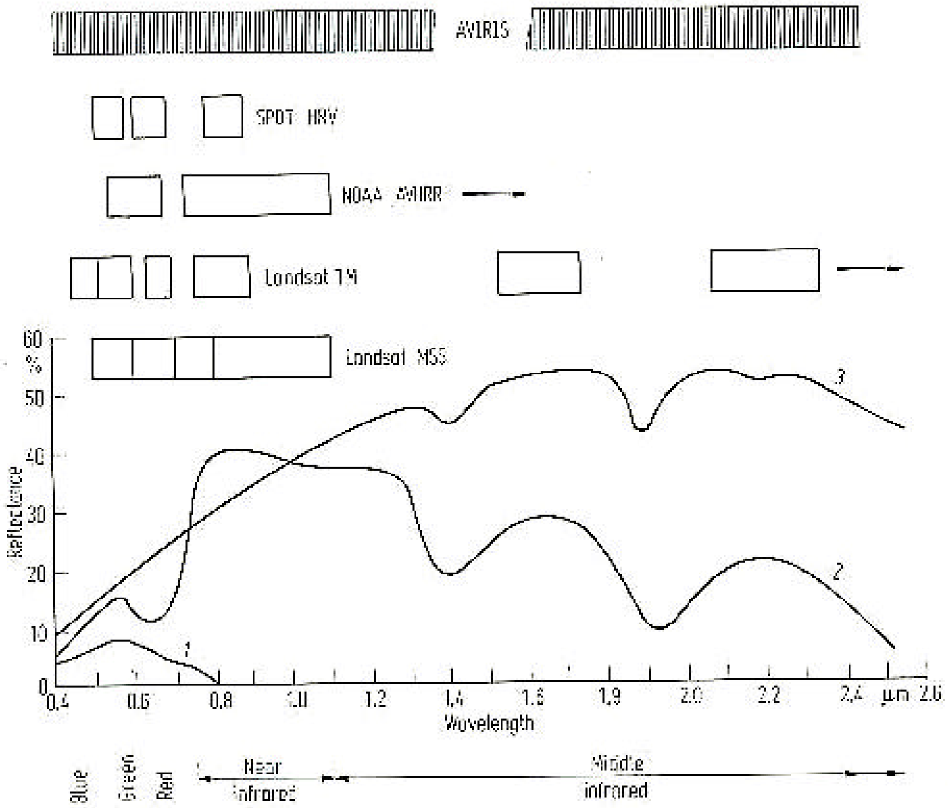
At the ***thermal end*** of the infrared range it is heat capacity and other thermal properties of the surface and near subsurface that control the strength of radiation detected.

In the ***microwave range***, using active imaging systems determine the magnitude of the reflected signal:
- radar techniques,
- the roughness of the cover type being detected,
- and its electrical properties, expressed in terms of complex permittivity (which in turn is strongly influenced by moisture content).

In the range ***20 – 60 GHz***, atmospheric oxygen and water vapour have a strong effect on transmission and thus can be inferred by measurements in that range.

Thus each range of wavelength has its own strengths in terms of the information it can contribute to the remote sensing process. Consequently systems available are optimised for and operate in particular spectral ranges, and provide data that complements that from other sensors.

Figure depicts how the three dominant earth surface materials of soil, vegetation and water reflect the sun's energy in the visible/infrared range of wavelengths.



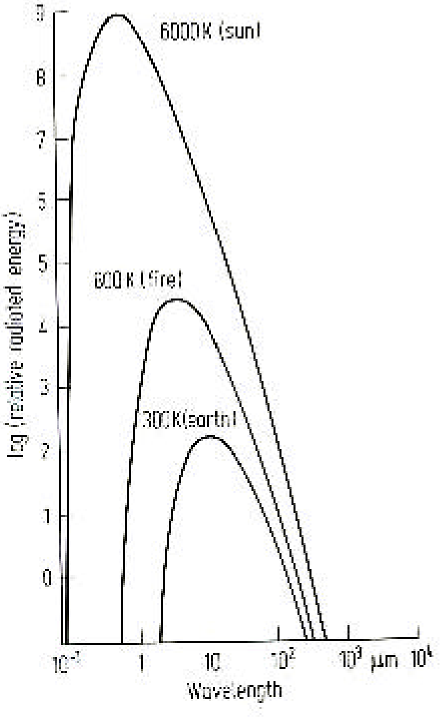
It is seen that water (graph 1) reflects about 10% or less in the blue-green range, a smaller percentage in the red and certainly no energy in the infrared range.
Should the water contain suspended sediments or should a clear water body be shallow enough to allow reflection from the bottom then an increase in apparent water reflection will occur, including a small but significant amount of energy in the near infrared range. This is a result of reflection from the suspension or bottom material.

Soils (graph3) have a reflection that increases approximately monotonically with wavelength, however with dips centred at about ***1.4 ìm***, ***1.9 ìm*** and ***2.7 ìm*** owing to moisture content. These water absorption bands are almost unnoticeable in very dry soils and sands. In addition to these bands clay soils also have hydroxyl absorption bands at ***1.4 ìm*** and ***2.2 ìm***.

The vegetation curve (graph 2) is considerably more complex than the other two. In the middle infrared range it is dominated by the water absorption bands at ***1.4 ìm***, ***1.9 ìm*** and ***2.7 ìm*** The plateau between about ***0.7 ìm*** and ***1.3 ìm*** is dominated by plant cell structure while in the visible range of wavelengths it is plant pigmentation that is the major determinant.

The curve sketched in figure is for ***healthy*** green vegetation. This has chlorophyll absorption bands in the blue and red regions leaving only green reflection of any significance. This is why we see chlorophyll pigmented plants as green.

In wavelength ranges between about ***3*** and ***14 ìm*** the level of solar energy actually irradiating the earth's surface is small owing to both the small amount of energy leaving the sun in this range by comparison to the higher levels in the visible and near infrared range (see figure below):



and the presence of strong atmospheric absorption bands between ***2.6 ìm*** and ***3.0 ìm, 4.2 ìm*** and ***4.4 ìm,*** and ***5 ìm*** and ***8 ìm***. Consequently much remote sensing in these bands is of energy being emitted from the earth's surface or objects on the ground rather than of 'reflected solar radiation.

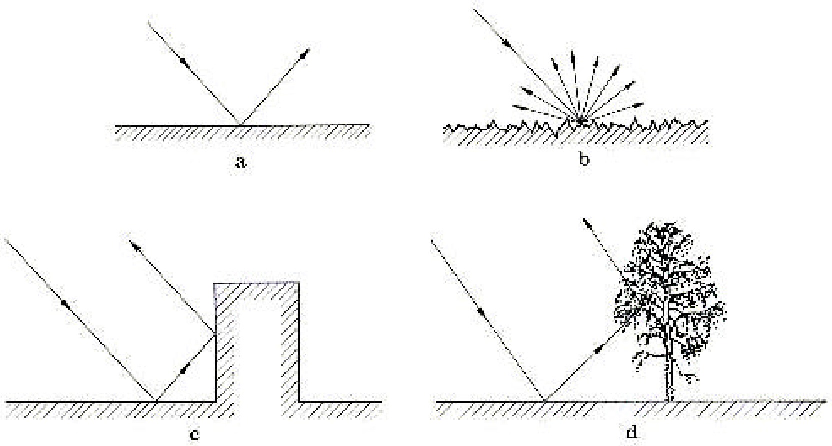
Figure above shows the relative amount of energy radiated from perfect black bodies of different temperatures. As seen, the sun at *6000 K* radiates maximally in the visible and near infrared regime but by comparison generates little radiation in the range around ***10 ìm***.

Incidentally, the figure shown does not take any account of how the level of solar radiation is dispersed through the inverse square law process in its travel from the sun to the earth. Consequently if it is desired to compare that curve to others corresponding to black bodies on the earth's surface then it should be considerably reduced overall.

The earth, at a temperature of about ***300 K*** has its maximum emission around ***10 ìm*** to ***12 ìm.*** Thus a sensor with sensitivity in this range will measure the amount of heat being radiated from the earth itself. Hot bodies on the earth's surface, such as bushfires, at around 800 K have a maximum emission in the range of about ***3 ìm*** to ***5 ìm***. Consequently to map fires, a sensor operating in that range would be used.

Real objects do not behave as perfect black body radiators but rather emit energy at a lower level than that shown in figure above. The degree to which an object radiates by comparison to a black body is referred to as its emittance. Thermal remote sensing is sensitive therefore to a combination of an object's temperature and emittance, the last being wavelength dependent.

Microwave remote sensing image data is gathered by measuring the strength of energy scattered back to the satellite or aircraft in response to energy transmitted. The degree of reflection is characterized by the scattering coefficient for the surface material being imaged. This is a function of the electrical complex permittivity of the material and the roughness of the surface in comparison to a wavelength of the radiation used.
Reflection in the direction of scattering may be away from the incident direction:



Smooth surfaces appear dark to black in image data. Rough surfaces act as diffuse reflectors; they scatter the incident energy in all directions, including back towards the remote sensing platform. As a result they appear light in image data. A third type of surface scattering mechanism is often encountered in microwave image data, particularly associated with man-made features such as buildings, and gives a very bright response.

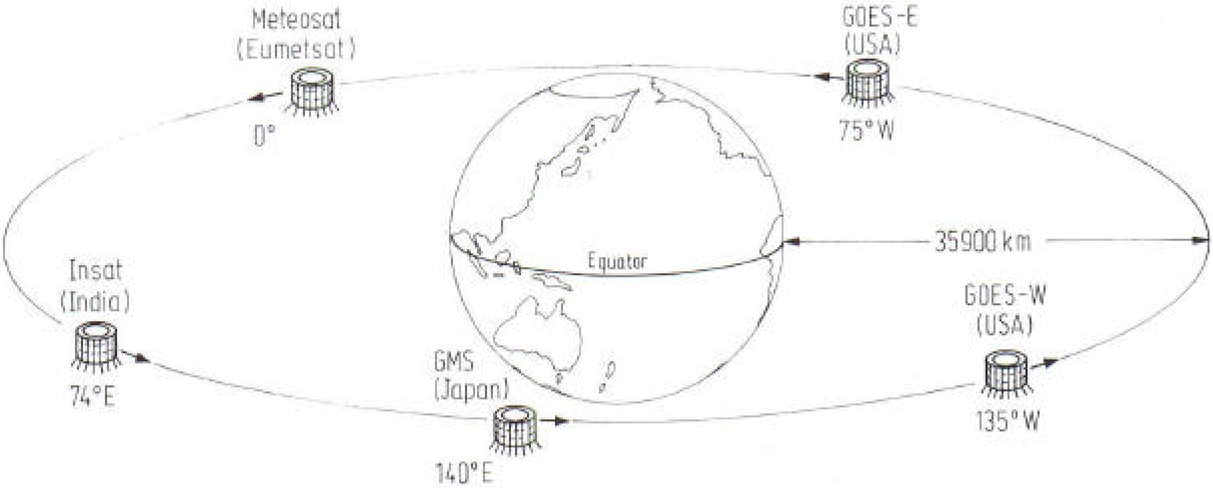
In interpreting image data acquired in the microwave region of the electromagnetic spectrum it is important to recognise that the four reflection mechanisms are present and modify substantially the tonal differences resulting from surface complex permittivity variations. By comparison, imaging in the visible/infrared range in which the sun is the energy source, results almost always in diffuse reflection, allowing the interpreter to concentrate on tonal variations resulting from factors such as soil, water, vegetation.

Weather Satellite Sensors

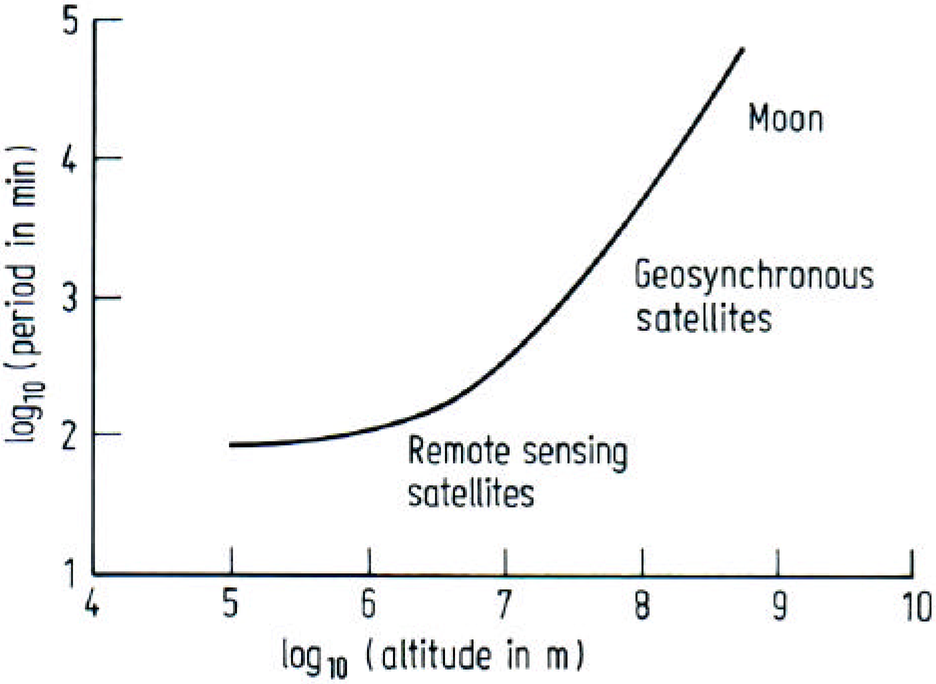
Weather satellites and those used for earth resources sensing operate in much the same bands of wavelength. Perhaps the major distinction in the image data they provide lies in the spatial resolutions available. Whereas data acquired for earth resources purposes generally has a pixel size less than ***100 m***, that used for meteorological applications usually has a much coarser pixel – often of the order of ***1 km x 1 km***. This is the distinction used herein in order to separate the two types of sensor. Having made that distinction however it is important to note that because of the similarity in wavebands, meteorological satellite data such as that from the NOAA Advanced Very High Resolution Radiometer (AVHRR) does find application in remote sensing when large synoptic views are required.

Polar Orbiting and Geosynchronous Satellites

Two broad types of weather satellite are in common use. One is of the polar orbiting, or more generally low earth orbit, variety whereas the other is at geosynchronous altitudes. The former typically have orbits at altitudes of about ***700 to 1500 km*** whereas the geostationary altitude is approximately ***36,000 km***.



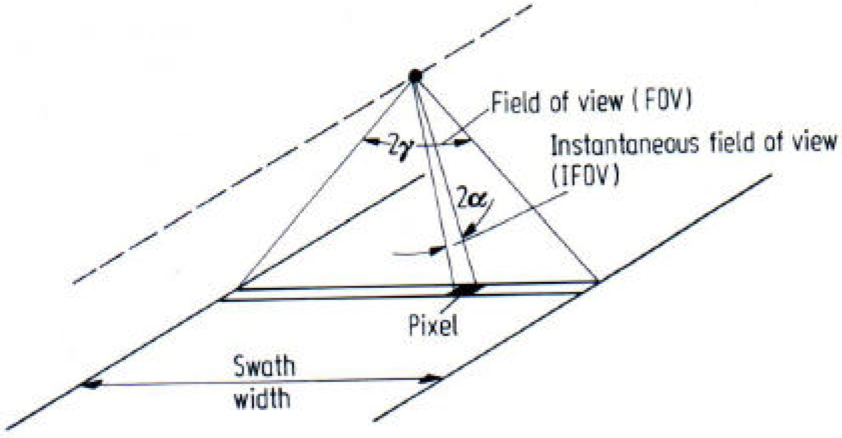
Satellite periods versus altitude:



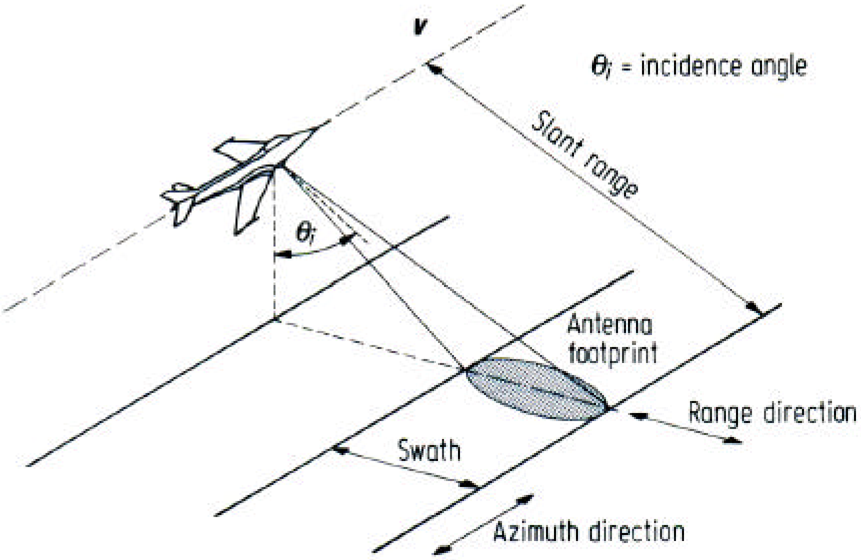
Characteristics of the radiometers include:

- ground resolution,
- dynamic range,
- ground swath,
- spectral bands (several).

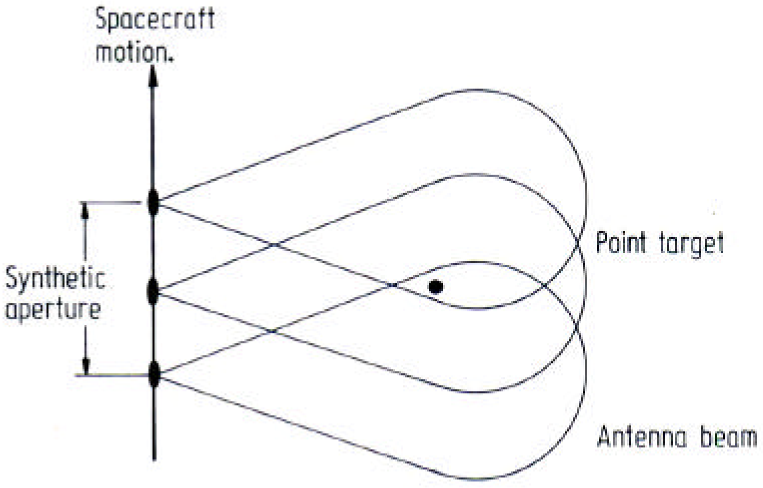
The concept of field of view (FOV) and instanteneous field of view (IFOV):



Side looking airborne radar:



Synthesis aperture radar (the concept of synthesizing a large antenna by utilyzing spacecraft motion along its orbital path):



Types of spatial data

As of sources of multispectral digital image data of the earth's surface; each image considered has represented the spatial distribution of energy coming from the earth in one or several wavelength ranges in the electromagnetic spectrum.

Other sources of spatially distributed data are also often available for regions of interest. These include simple maps that show topography, land ownship, roads and the like, through to more specialised sources of spatial data such as maps of geophysical measurements of the area.

Frequently these other spatial data sources contain information not available in multispectral imagery and often judicious combinations of multispectral and other map-like data allow inferences to be drawn about regions on the earth's surface not possible when using a single source on its own.

Consequently the image analyst ought to be aware of the range of spatial data available for a region and select that subset likely to assist in the information extraction process.

Table below is an illustration of the range of spatial data one might expect could be available for a given region. This differentiates the data into three types according as to whether it represents point information, line information or area information.

Table: Sources of spatial data

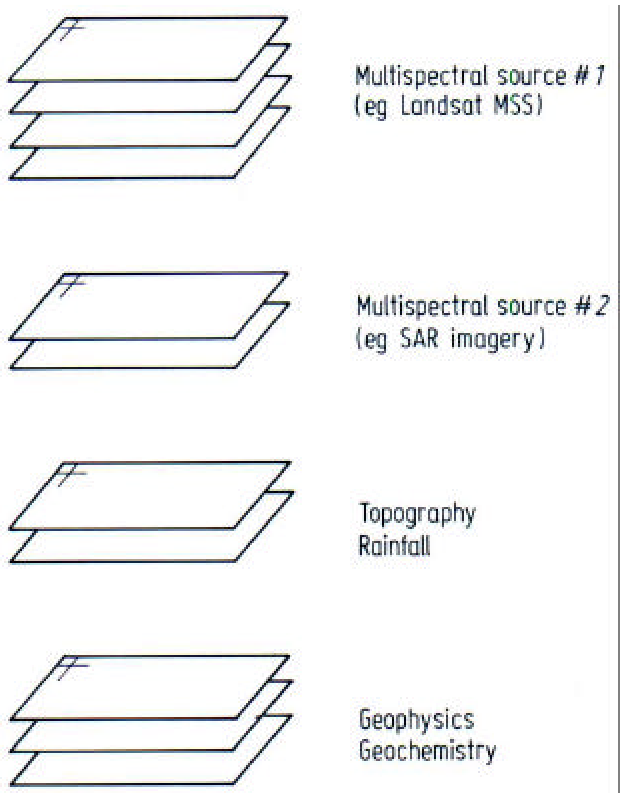
Point	Line	Area
Multispectral data	road maps	land ownership
Topography	powerline grids	town plans
Magnetic measurements	pipeline networks	geological maps
Gravity measurements		land use licenses
Radiometric measurements		land use maps
Rainfall		land cover maps
Geochemistry (in ppm)		soil type maps

Irrespective of type however, for a spatial data set to be manipulated using the techniques of digital image processing it must share two characteristics with multispectral data of the types:

1. it must be available in discrete form spatially, and in value (so, it must consist of, or be able to be converted to, pixels with each pixel describing the properties of a given (small) area on the ground: the value ascribed to each pixel must be expressible in digital form)
2. it must be in correct geographic relation to a multispectral image data set if the two are to be manipulated together.

In situations where multispectral data is not used, the pixels in the spatial data source would normally be arranged to be referenced to a map grid. It is usual however, in digital spatial data handling systems, to have all entries in the data set relating to a particular geographical region, mutually registered and referenced to a map base (such as the UTM grid system).

When available in this manner the **data is said to be geocoded**. Means by which different data sets can be registered are to be introduced later. Such a database is depicted in the figure:



Data Formats

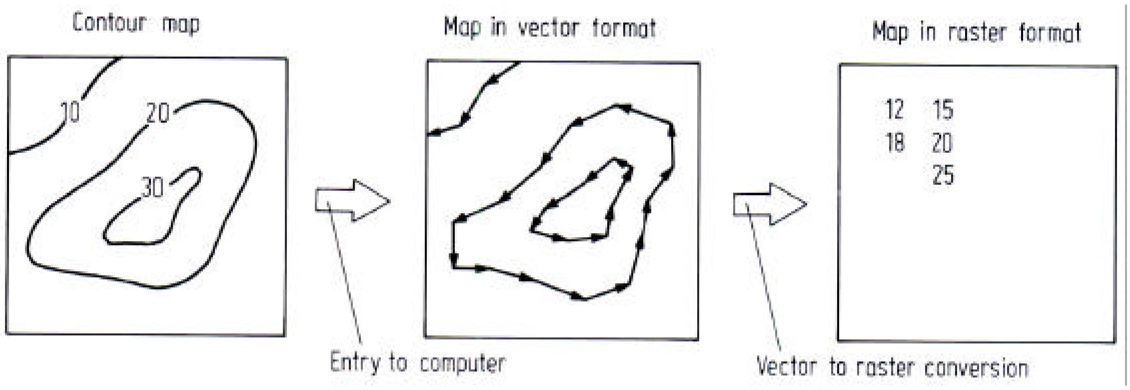
Not all sources of spatial data are available originally in the pixel oriented digital format.

Indeed often the data will be available as **analog maps** that require digitisation before entry into a digital data base. That is particularly the case with ***line and area*** data types, in which case also consideration has to be given to the "value" that will be ascribed to a particular pixel.

In ***line*** spatial data sources the pixels could be called zero if they were not part of a line and coded to some other number if they formed part of a line of a given type. For a road map, for example, pixels that fall on highways might be given a value of I whereas those on secondary roads could be given a value of 2, and so on. On display, the different numbers could be interpreted and output as different colours.

In a similar manner numbers can be assigned to different regions when digitizing ***area spatial data*** sources.

Conceptually the digitization process may ***not be straightforward***. Consider the case for example of needing to create a digital topographic map from its analog contour map counterpart. Figure below illustrates this process:



First it is necessary to convert the contours on the paper map to records contained in a computer. This is done by using an input device such as a ***stylus or cross-hair*** cursor to mark a series of points on each contour between which the contour is regarded by the computer to be a straight line.

Information on a contour at this stage is stored in the computer's memory as a **file** of points. This format is referred to as ***vector format*** owing to the vectors that can be drawn from point to point (in principle) to reconstruct a contour on a display or ***plotter*** (some spatial data handling computer systems operate in vector format entirely).

However to be able to exploit the techniques of digital image processing the vector formated data has to be turned into a set of pixels arranged on rectangular grid centres. This is referred to as *rasterformat* (or sometimes grid format); the elevation values for each pixel in the raster form are obtained by a process of interpolation over the points recorded on the contours. The operation is referred to as *vector to raster conversion* and is an essential step in entering map data into a digital spatial data base.

Raster format is a natural one for the representation of multispectral image data since data of that type is generated by digitising scanners, is transmitted digitally and is recorded digitally. Moreover many ***image forming devices*** such as filmwriters and television monitors operate on a raster display basis, compatible with digital data acquisition and storage.

Raster format however is also appealing from ***a processing point of view*** since the logical records for the data are the pixel values (irrespective of whether the data is of the point, line or area type) and neighbourhood relationships are easy to establish by means of the pixel addresses.

This is important for processing operations that involve ***near neighbouring groups*** of pixels. In contrast, vector format ***does not offer*** this feature. However an advantage of vector format, often exploited in ***high quality graphics*** display devices, is that resolution is not limited by pixel size.

Geographic Information Systems (GIS)

The amount of data to be handled in a database that contains spatial sources such as satellite and aircraft imagery along with maps, ***is enormous***, particularly if the data covers a large geographical region. Quite clearly therefore thought has to be given to ***efficient means*** by which the data types can be stored and retrieved, manipulated, analysed and displayed.

This is the role of the geographic information system (GIS). Like its commercial counterpart, the management information system (MIS), the GIS is designed to carry out operations on the data stored in its database, according to a set of ***user specifications***, without the user needing to be knowledgeable about how the data is stored and what data handling and processing procedures are utilized to retrieve and present the information required.

Unfortunately because of the nature and volume of data involved in a GIS ***many of*** the MIS concepts developed for data base management systems (DBMS) cannot be transferred directly to GIS design although they do provide guidelines. Instead ***new design concepts*** have been needed, incorporating the sorts of operation normally carried out with spatial data, and attention has had to be given to efficient coding techniques to facilitate searching through the large numbers of maps and images often involved.

To understand the sorts of spatial data manipulation operations of importance in GIS one must take the view of the resource manager rather than the data analyst. Whereas the latter is concerned with image reconstruction, filtering, transformation and classification, the manager is interested in operations such as those listed in table:

Table. Some GIS data manipulation operations
• Intersection and overlay of data sets (masking)
• Intersection and overlay of polygons with spatial data identification of shapes
• Identification of points in polygons
• Area determination
• Distance determination
• Thematic mapping
• Proximity calculations (shortest route, etc.)
• Search by data
• Search by location
• Search by user-defined attribute
• Similarity searching (e.g. of images)

These provide information from which management strategies and the like can be inferred. Certainly, to be able to implement many, if not most, of these a ***substantial amount of image processing*** may be required.

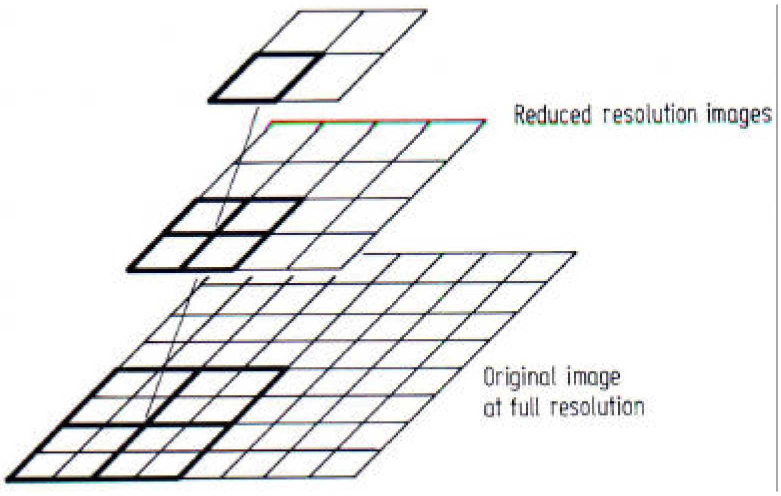
A problem which can arise in image data bases of the type encountered in a GIS is the need to identify one image by reason of its ***similarity*** to another. In principle, this could be done by comparing the images *pixel-by-pixel*; however the computational demand in so doing would be *enormous* for images of any practical size.

Instead effort has been directed to developing ***codes or signatures*** for complete images that will allow efficient similarity searching.

For example ***an image histogram*** could be used, however as geometric detail is not preserved in a histogram this is rarely a suitable code for an image on its own.

One ***effective possibility*** that has been explored is the use of image pyramids. A pyramid is created by *combining groups of pixels in a neighbourhood* to produce a new composite pixel of reduced resolution, and thus a low resolution image with fewer pixels.

This process is *repeated* on the processed image to form a new image of lower resolution (and fewer pixels) still. Ultimately the image could be reduced to one single pixel that is *a global measure of the image's brightness.* Since pixels are combined in neighbourhood groups, spatial detail is propagated up through the pyramid, albeit at decreasing resolution:



It is a relatively easy matter to show that the additional memory required to store a complete pyramid, constructed as in the figure, is only 33% more than that required to store just the image itself.

Having developed an image pyramid, ***signatures*** that can be used to undertake similarity searching include the histograms computed over rows and colums in the uppermost levels of the pyramid. A little thought shows that this allows an *enormous number of images to be addressed*, particularly if each pixel is represented by an 8 bit brightness value. As a result very fast searching can be carried out on these reduced representations of images.

There is sometimes an *image processing advantage* to be obtained when using a pyramid representation of an image. In ***edge detection***, for example, it is possible to localise edges ***quickly***, without having to search every pixel of an image, by finding apparent edges (regions) in the upper levels of the pyramid.

The succeeding lower pixel groupings are then searched to localise the edges better.

Finally the *pyramid representation* of an image is felt to have some *relation* to human perception of images. The ***upper levels contain global features*** and are therefore not unlike the picture we have when first looking at a scene generally we take the scene in initially "as a whole" and either miss or ignore detail.

Then we focus on regions of interest for which we pay attention to detail because of the information it provides us with.

The Challenge to Image Processing and Analysis

Much of the experience gained with ***digital image processing and analysis*** in remote sensing has been with multispectral image data.

Information extraction from geophysical data could be facilitated, for example, if a degree of sharpening is applied prior to photointerpretation, while colour density slicing could assist the interpretation of topography.

However the real challenge to the image analyst arises when data of mixed types are to be processed together.

The first relates to differences in resolution, an issue that arises also when treating multi-source satellite data such as Landsat MSS and NOAA AVHRR. The analyst must decide, for example what common pixel size will be used when co-registering the data, since either resolution or coverage will normally be sacrificed.

Clearly this decision will be based on the needs of a particular application and is a challenge more to the analyst than the algorithms.

The more important consideration however is in relation to techniques for machine assisted interpretation. There is little doubt that combined multispectral and, say, topographic or land ownership maps can yield more precise thematic (i.e. category of land cover, etc.) information for a particular region than the multispectral data on its own.

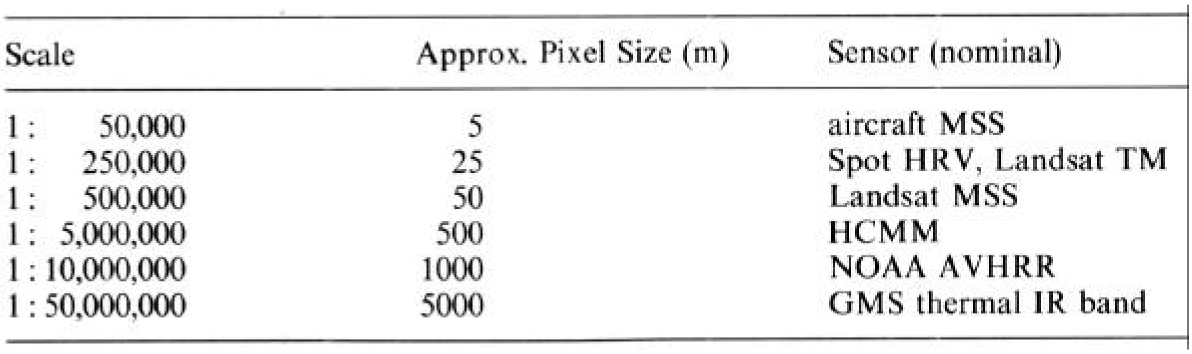
Indeed the combination of these sources is often employed in photointerpretive studies. However digitally, acceptable algorithms have yet to be devised that will permit integrated but diverse data types to be analysed automatically as effectively, and with the theoretical foundations, as have satellite and aircraft spectral data.

Such a basis will be required if geographic information systems are to incorporate a significant degree of digital processing and analysis between data retrieval and display.

The issue is complicated further when it is recalled that much of the non-spectral, spatial data available is not in numerical point form but rather is in nominal area or line fortnat. With these it seems clear that ***image analysis algorithms developed algebraically*** will not be suitable. Rather some degree of logical processing of labels combined with algebraic processing of arithmetic values (such as pixel brightnesses) will be necessary.

A Comparison of Scales in Digital Image Data

Because of IFOV differences the digital images provided by various remote sensing sensors will find application at different scales. Thus Landsat MSS data is suggested as being suitable for scales smaller than about 1 :500,000 whereas NOAA AVHRR data is suitable for scales below 1 : 10,000,000.



Scale	Approx. Pixel Size (m)	Sensor (nominal)
1 : 50,000	5	aircraft MSS
1 : 250,000	25	Spot HRV, Landsat TM
1 : 500,000	50	Landsat MSS
1 : 5,000,000	500	HCMM
1 : 10,000,000	1000	NOAA AVHRR
1 : 50,000,000	5000	GMS thermal IR band