Fundamentals of GIS and Remote Sensing

Fundamentals of GIS
and
Remote Sensing

Dr. Udhav Eknath Chavan

GARIMA PRAKASHAN
KANPUR-208 021 (INDIA)

Fundamentals of GIS and Remote Sensing

First Published : 2019
ISBN : 978-93-82730-67-5

Published by

GARIMA PRAKASHAN
1st Floor, Mayur Park, Basant Vihar
Naubasta, Kanpur - 21
Mob. : 09415200584,9506294444, 09415125867
Telefax-0512-2634444(P.P.)
Email : garima.prakashan@gmail.com
Website-www.garimaprakashan.com

PRINTED IN INDIA

Printed at Deepak Offset Press, Delhi.

Preface

Geographical Information System (GIS) techniques are being widely used since last two decades, mainly by earth scientists, geologists, geographers, regional planners, environmentalists in the fields of mineral explorations, archaeological excavations, geographical and topographical analysis, terrain studies, urban planning, regional planning, natural resources management, natural disaster management etc.

A geographic information system, geographical information system, or geospatial information system is any system that captures, stores, analyses, manages, and presents data that are linked to location. In the simplest terms, GIS is the merging of cartography, statistical analysis, and database technology. GIS may be used in archaeology, geography, cartography, remote sensing, land surveying, public utility management, natural resource management, precision agriculture, photogrammetry, urban planning, emergency management, landscape architecture, navigation, aerial video, and localized search engines.

Remote sensing is nothing but a means to get the reliable information about an object without being in physical contact with the object. It is on the observation of an object by a device separated from it by some distance utilizing the characteristics response of different objects to emissions in the electromagnetic energy is measured in a number of spectral bands for the purpose of identification of the object.

Remote sensing may be taken to mean the observation of, or gathering information about, a target by a device separated from it by some distance. Remote sensing instruments are of two primary types—active and passive. Active sensors, provide their own source of energy to illuminate the objects they observe. In satellite remote sensing of the earth, the sensors are looking through a layer of atmosphere separating the sensors from the Earth's surface being observed.

In Optical Remote Sensing, optical sensors detect solar radiation reflected or scattered from the earth, forming images resembling photographs taken by a camera high up in space. The wavelength region usually extends from the visible and near infrared (commonly abbreviated as VNIR) to the short-wave infrared (SWIR). Different materials such as water, soil, vegetation, buildings and roads reflect visible and infrared light in different ways. They have different colours and brightness when seen under the sun. The interpretation of optical images require the knowledge of the spectral reflectance signatures of the various materials (natural or man-made) covering the surface of the earth. There are also infrared sensors measuring the thermal infrared radiation emitted from the earth, from which the land or sea surface temperature can be derived.

Photogrammetry is the practice of determining the geometric properties of objects from photographic images. Photogrammetry is as old as modern photography and can be dated to the mid-nineteenth century. In the simplest example, the distance between two points that lie on a plane parallel to the '. Photogrammetry is used in different fields, such as topographic mapping, architecture, engineering, manufacturing, quality control, police investigation, angy]], as well as by archaeologists to quickly produce plans of large or complex sites and by meteorologists as a way to determine the actual wind speed of a tornado where objective weather data cannot be obtained. It is also used to combine live action with computer generated imagery in movie post-production; *Fight Club* is a good example of the use of photogrammetry in film.

The book provides a comprehensive coverage of techniques, applications and technologies of remote sensing and Geographical Information System (GIS). . This book should be useful to remote sensing scientists and engineers, geographers, geologists, ecologists and environmental scientists, agricultural and soil scientists.

—*Dr. Udhav Eknath Chavan*

Contents

1

Introduction to GIS

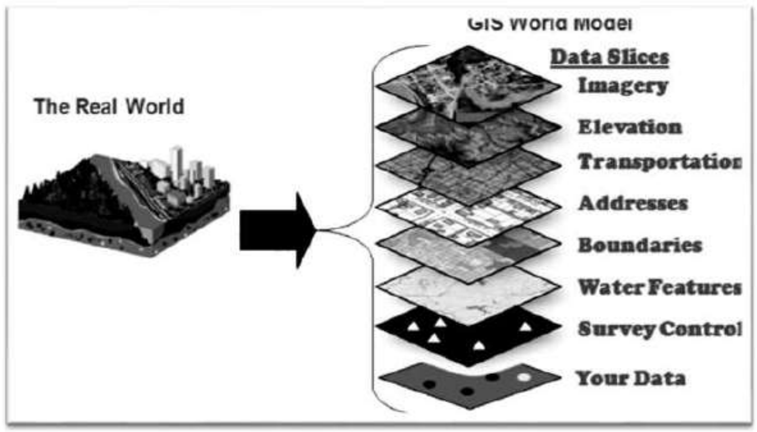
DEFINITION, DEVELOPMENT, COMPONENTS, OBJECTIVES, HARDWARE AND SOFTWARE REQUIREMENTS

Definition

Geographic Information System (GIS) is a computer system build to capture, store, manipulate, analyze, manage and display all kinds of spatial or geographical data. GIS application are tools that allow end users to perform spatial query, analysis, edit spatial data and create hard copy maps. In simple way GIS can be define as an image that is referenced to the earth or has x and y coordinate and it's attribute values are stored in the table. These x and y coordinates are based on different projection system and there are various types of projection system. Most of the time GIS is used to create maps and to print. To perform the basic task in GIS, layers are combined, edited and designed.

GIS can be used to solve the location based question such as "What is located here" or Where to find particular features? GIS User can retrieve the value from the map, such as how much is the forest area on the land use map. This is done using the query builder tool. Next important features of the GIS is the capability to combine different layers to show new information. For example, you can combine elevation data, river data, land use data and many more to show information about the landscape of the area.

From map you can tell where is high lands or where is the best place to build house, which has the river view . GIS helps to find new information.



Gis Model

How GIS Works:

- Visualizing Data: The geographic data that is stored in the databases are displayed in the GIS software.
- Combining Data: Layers are combined to form a maps of desire.
- The Query: To search the value in the layer or making a geographic queries.

Definition by others: A geographic information system (GIS) lets us visualize, question, analyze, and interpret data to understand relationships, patterns, and trends. (ESRI)

Component of GIS

Hardware: Hardware is the physical component of the computer and GIS runs on it. Hardware may be hard disk, processor, motherboard and so on. All these hardware work together to function as a computer. GIS software run on these hardware. Computer can be standalone called desktop or server based. GIS can run on both of them.

Software: GIS Software provides tools and functions to input and store spatial data or geographic data. It provides tool to perform geographic query, run analysis model and display geographic data in the map form. GIS software uses Relation Database Management System (RDBMS) to store the geographic data. Software talks with the database to perform geographic query.

Data: Data are the fuel for the GIS and the most important and expensive component. Geographic data are the combination of physical features and it's information which is stored in the tables. These tables are maintained by the RDBMS. The process of capturing the geographic data are called digitization which is the most tedious job. It is the process of converting scanned hardcopy maps into the digital format. Digitization is done by tracing the lines along the geographic features for example to capture a building you will trace around the building on the image.

People: People are the user of the GIS system.

People use all above three component to run a GIS system. Today's computer are fast and user friendly which makes it easy to perform geographic queries, analysis and displaying maps. Today everybody uses GIS to perform their daily job.

History And Development Of GIS

The recent historical narrative of geographical information systems is no less interesting than the systems themselves as they materialized rather quickly over a relatively short span of time.

This can be explained in phases:

PHASE 1 – Beginning in the early 1960s and lasting almost 15 years, this is when geographical information systems took flight and became noticed all around the World. GIS at this time was reserved to a select few organizations with deep pockets.

PHASE 2 – In the late 1970's Geographic Information Systems captured National attention, often being used in Government agencies. The true value of these GIS and it's ability to analyze and interpret large ammounts of data like no other system before was realized.

PHASE 3 – For the next decade, geographical information systems were marketed to the general public. Smaller organizations now had access to these complex information systems, and the popularity spread quickly.

Coming under the fold of market-based commercialization, which allows a greater number of people and entities to interact with it. In these years, GIS technology was tailored to everyday users instead of only technical masterminds.

PHASE 4 – Since the mid-90's GIS technology has continuously grown with new inventions and ground-breaking advancements. Today, we use Geographic Information Systems everyday in multiple industries and countries around the world.

GIS Objectives

- Maximize the efficiency of decision making and planning.
- Provide efficient means for data distribution and handling.
- Elimination of redundant database-minimize duplication.
- Capacity to integrate information from many sources.
- Complex analysis/queries involving geographical reference data to generate new information.
- Update data quickly and cheaply.

For any application there are five generic questions a GIS can answer-

- Location: What exists at a particular location? This question seeks to find out for the answer like, location of a particular object or area in terms of latitude/longitude or X/Y.
- Condition: Identify where certain condition exists. This tends to answer for all those questions where certain conditions are satisfied.
- Trends: What has changed since? This question is applied to find a noticeable difference or change incurred within a particular time period.

- Pattern: what spatial pattern exists? This is the most logical question is answered by GIS that distribution of spatial features and reasons behind that distribution.

Hardware and Software Requirements for GIS

Many GIS applications are implemented within organizations and have to play an important role. The effectiveness of GIS projects therefore depends as much on the futuristic planning of hardware and software requirements, as on technicalities of its implementation. It is therefore necessary to address the question of how GIS can be successful in an organization and continue to do so for many years before upgradations of hardware and software are needed. GIS projects are successful because they help the organisation to perform better in society. But it is also true that organisational issues will contribute to the long-term success of a GIS project. In other words, GIS and the host organisation have a close and two-way relationship.

It is neither possible nor desirable to recommend here a generalized requirement in terms of hardware and software for a GIS program. Nevertheless, there is a necessity to analyse the requirements keeping in view the following:

1. Organization's goals and objectives
2. Analysis of current practices and problems encountered and redundancies
3. Organizational issues - management structure, resources, staff
4. training
5. Applications
6. Base map requirements
7. Conceptual data base design
8. System integration
9. Cost of GIS implementation – applications development, data conversion, hardware, software, resources, training, update and maintenance

10. The Strategic Implementation Plan over a period of 3-5 years

Hardware Requirement

It is obvious that hardware and software requirements vary considerably depending on the tasks undertaken. The following minimum configuration allows installation of most modern GIS applications for work with small components. Recommended configurations are noted in parentheses for work with anything other than small drawings.

- 2.8 Ghz PIV true PC compatible (dual-core processor recommended).
- 1 GB RAM (4 GB or greater recommended).
- 800 x 600 SVGA Display (1280 x 1024 or greater recommended).
- 250 MB hard disk free space (gigabytes of free space recommended).
- Windows 2000, Windows XP, Windows Server 2003, Windows Server 2008 or Vista with most recent service pack, in standard 32-bit versions or in 64-bit versions. (Windows XP or greater recommended).
- Internet Explorer 6 or most recent IE version plus most recent service pack.
- Microsoft's .NET Framework 2.0 or more recent.
- IIS 5.1 or greater to operate IMS.

For very large tasks, such as intensive web server applications using IMS, investing in a dual socket, quad processor machine may be considered. This has eight cores. Various processes within modern GIS applications can use multiple processors.

Memory Requirements for Large Projects - Manifold is designed for an era where RAM memory is cheap and personnel costs are high. Undo and many other user-friendly features require a lot of memory to implement. For best performance we recommend installing lots of RAM and having plenty of free disk space for temporary files. RAM is so cheap in modern times it is "penny wise and pound foolish" not to load up your system with the maximum amount of RAM it can hold. This will help all your work go faster and not just Manifold. As always, keep in mind that to really use lots of RAM effectively you should be running 64-bit Windows such as Windows Vista x64, Windows Server 2008 x64 or Windows XP x64.

Software Requirement

The following is a list of GIS application packages from which one can choose. The choice shall depend on the needs of the organization, functionality desired and the money available, and the period for which the planning is being done. One may need to make a comparison of costs and benefits (both of which keep changing rapidly) before making a final decision.

THE BASIS OF GIS MAPPING: MAP PROJECTIONS, DATUM AND COORDINATE SYSTEMS

A Geographic Information System (GIS) is a software package that is helping to digitize the world around us. Digital maps and location based information helps government agencies, businesses and even people like you and me keep up with the changing pace of the world around us. GIS applications improve efficiencies, reduce costs, and bring together data in ways that wasn't possible even 10 years ago. This chapter focuses on some of the basic components and aspects of Geographic Information Systems.

Geographically Based Data Sets

A Geographic Information System (GIS) is a system of computer applications that can be used to display, manipulate, and analyze spatially varied information from multiple sources all in one place. Most often the datasets used in a GIS are categorized into multiple categories for easier storage and use. Each dataset that a GIS can support is divided into two main parts: graphical (spatial) information and tabular (attribute) information. Spatial data is data that is geo-referenced or location specific and is what is shown graphically on the computer screen. Each piece of graphic information is called a feature. Features can be points, lined or even polygons.

The attribute or tabular information is text based or numerical information that describe each of the features. The tabular information is linked to the graphical information and includes a unique ID number used to represent each point, line or polygon. Examples of tabular data can include such things as addresses, coordinates, area, length, sales information, road names, etc. The possibilities for data association between tabular and graphical information are endless.

Spatial data can often be further divided into two major types. The first type is raster data which is usually in the form of images such as aerial photographs or imported scans of old maps. The raster data stores the location and color value of each pixel that forms the image. The second type of data is vector based. Vectors can be a combination of linework, polygons and curvilinear data. This information is stored using a combination of location specific point, lines, and arcs. Raster imagery can lose quality and become blurred when scaled. However, vector data is scalable to any size without losing any integrity.

Another way to look at data types is understanding that some data is discrete while others are continuous. Discrete data is usually vector based and has specific information located at specific points with gaps in between. On the other hand continuous data is usually raster based and no gaps are present. Anywhere within the domain of a raster there will be information.

When each dataset is loaded and displayed in the GIS map window it is called a layer. When multiple layers are used in the same map window they can be stacked, color coded, and symbolized to represent an endless array of map compositions. Stacked datasets can also be manipulated by adjusting the colors ramps, hatching, shading and/or transparency levels to reveal new relationships that would not have otherwise been obvious with traditional maps. In essence, the main purpose of a GIS is to describe, analyze and display a variety of spatial data in a way that only a digital map can.

Components of a GIS Software Application

On the surface a GIS is simply a combination of software, hardware and data. However, more specifically, a GIS usually has a data hub (typically a collection of data on a server) and a graphical user interface for the people who will be using and manipulating the data. In this way all users of the GIS can be connected to the same data sets. When data sets are updated or augmented everyone who uses the data can see the changes. This ensures consistency of information and will also help to avoid duplication of work. Web-based applications can also be added so that remote users can see the same information that everyone else can.

Differences between CAD and GIS

In some ways CAD and GIS software are very similar. Both software tools can display data such as points, lines, and polygons to a specific scale. Both CAD and GIS tools can also be used to create maps and representative drawings. The similarities between the two types of software applications usually ends there. Typically CAD is used for design of intricately detailed objects such as manufactured products, buildings, commercial site layouts and roadways. The common theme with CAD work is that it is used for representing manmade projects. However GIS is used to analyze information on a much larger scale and can often encompass things that make up the natural world such as forests, soil strata and even rivers and floodplains. As mentioned previously, GIS is great for attaching tabular data to graphical information while most CAD tools lack that ability. CAD is a great engineering tool however its capabilities are not honed for most geographically based work that a GIS can perform.

Differences Between GIS and Spreadsheets/Databases

The main difference between spreadsheets/databases and a GIS is that a GIS uses geography and location information as the main piece of information to display and relate data with other data. Even a relational spreadsheet or complex database lacks the clear functionality of a graphical tool that bases everything on location. When it comes to location information spreadsheets limited to simply listing numerical information such as coordinates (latitude, longitude, etc) or addresses in rows and columns. A spreadsheet with this kind of information can easily be converted into a form that can be imported to a GIS. In addition to this, data within a GIS can have spreadsheet like structures supporting them. For example, you could have a series of lines representing a roadway network with a spreadsheet joined to those lines with information such as number of lanes, roadway width, pavement surface, etcetera for each line segment.

Benefits of Geographically Based Data Integration

By far the most obvious benefit of creating a GIS is the fact that it eliminates all manual forms of geography based analysis. There is no longer a need to print large maps on transparency sheets nor does one need a light table to get the effect of layering multiple printed sheets on top of one another. The affects and problems of differently scaled maps can also be addressed as a GIS can convert and project all data on the same coordinate system and scale. The second benefit of a GIS is that it centralizes and organizes the data for use in one comprehensive system. No longer will there be lost records and all of the data created and collected by the user can be stored and used again in the future with ease. And yet a third benefit is the fact that there is an infinite variety of maps can be created with even just a few datasets. Cartographers are no longer limited to present geographic information in only one way. Most GIS programs can easily symbolize or represent data in any manner imaginable and at essentially any scale.

Coordinate Systems

In order to represent spatially varied data everything needs to be placed on a common coordinate system. In the mapping world there are three main types of coordinate systems. The first system is called the Cartesian Coordinate system and can be represented by a grid with a numbering system that can locate information on a horizontal and vertical axis. The second systems is the polar coordinate system. This is an easy way to locate information about a central using only an angle and a distance (radius). In many junior high and high school math classes students learn about these basic coordinate systems. The third type of coordinate system is a global coordinate system. At its most basic level a global coordinate system is where two numbers (latitude and longitude) are used to reference a specific location on the earth.

Mapping Concepts, Features & Properties

A map represents geographic features or other spatial phenomena by graphically conveying information about locations and attributes.

Locational information describes the position of particular geographic features on the Earth's surface, as well as the spatial relationship between features, such as the shortest path from a fire station to a library, the proximity of competing businesses, and so on. Attribute information describes characteristics of the geographic features represented, such as the feature type, its name or number and quantitative information such as its area or length.

Thus the basic objective of mapping is to provide:

- descriptions of geographic phenomenon
- spatial and non spatial information
- map features like Point, Line, & Polygon.

Map Features

Locational information is usually represented by points for features such as wells and telephone pole locations, lines for features such as streams, pipelines and contour lines and areas for features such as lakes, counties and census tracts.

Point Feature

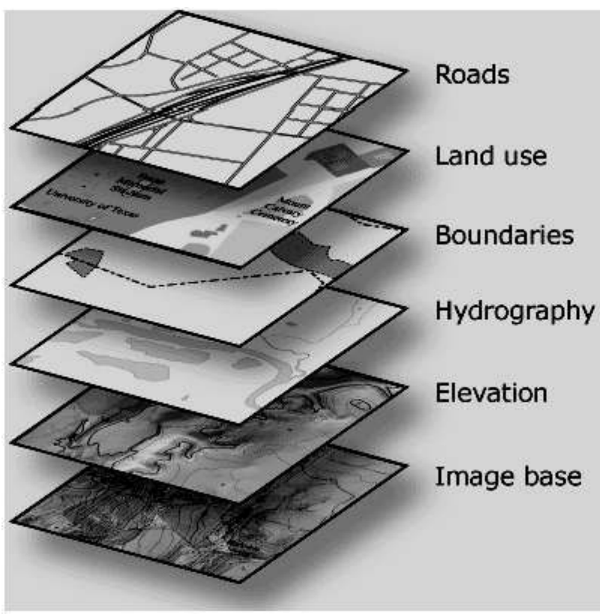
A point feature represents as single location. It defines a map object too small to show as a line or area feature. A special symbol of label usually depicts a point location.

Line Feature

A line feature is a set of connected, ordered coordinates representing the linear shape of a map object that may be too narrow to display as an area such as a road or feature with no width such as a contour line.

Map projections

Within ArcGIS, every dataset has a coordinate system, which is used to integrate it with other geographic data layers within a common coordinate framework such as a map. Coordinate systems enable you to integrate datasets within maps as well as to perform various integrated analytical operations such as overlaying data layers from disparate sources and coordinate systems.



Coordinate system

Coordinate systems enable geographic datasets to use common locations for integration.

A coordinate system is a reference system used to represent the locations of geographic features, imagery, and observations such as GPS locations within a common geographic framework.

Each coordinate system is defined by:

- Its measurement framework which is either geographic (in which spherical coordinates are measured from the earth's center) or planimetric (in which the earth's coordinates are projected onto a two-dimensional planar surface).
- Unit of measurement (typically feet or meters for projected coordinate systems or decimal degrees for latitude–longitude).
- The definition of the map projection for projected coordinate systems.
- Other measurement system properties such as a spheroid of reference, a datum, and projection parameters like one or more standard parallels, a central meridian, and possible shifts in the x- and y-directions.

Types of coordinate systems

There are two common types of coordinate systems used in GIS:

- A global or spherical coordinate system such as latitude–longitude. These are often referred to as*geographic coordinate systems.*
- A projected coordinate system based on a map projection such as transverse Mercator, Albers equal area, or Robinson, all of which (along with numerous other map projection models) provide various mechanisms to project maps of the earth's spherical surface onto a two-dimensional Cartesian coordinate plane. Projected coordinate systems are sometimes referred to as *map projections*.

Coordinate systems (either geographic or projected) provide a framework for defining real-world locations. In ArcGIS, the coordinate system is used as the method to automatically integrate the geographic locations from different datasets into a common coordinate framework for display and analysis.

ArcGIS automatically integrates datasets whose coordinate systems are known

All geographic datasets used in ArcGIS are assumed to have a well-defined coordinate system that enables them to be located in relation to the earth's surface.

If your datasets have a well-defined coordinate system, then ArcGIS can automatically integrate your datasets with others by projecting your data on the fly into the appropriate framework—for mapping, 3D visualization, analysis, and so forth.

If your datasets do not have a spatial reference, they cannot be easily integrated. You need to define one before you can use your data effectively in ArcGIS. The spatial reference or coordinate system is metadata. It describes the coordinate framework that the data is already using.

Caution: When you define the coordinate system for a dataset using the Define Projection tool or the dataset property page, you are updating the metadata to identify the current coordinate system.

The dataset's extent and coordinate values will not change. The dataset must already be using the coordinate system. To change a dataset's coordinate system, including its extent and values, use the Project or Project Raster tools.

Spatial reference in ArcGIS

A *spatial reference* in ArcGIS is a series of parameters that define the coordinate system and other spatial properties for each dataset in the geodatabase. It is typical that all datasets for the same area (and in the same geodatabase) use a common spatial reference definition.

An ArcGIS spatial reference includes settings for:

- The coordinate system
- The coordinate precision with which coordinates are stored (often referred to as the coordinate resolution)
- Processing tolerances (such as the cluster tolerance)
- The spatial or map extent covered by the dataset (often referred to as the spatial domain)

Projection Types

Because maps are flat, some of the simplest projections are made onto geometric shapes that can be flattened without stretching their surfaces. These are called developable surfaces. Some common examples are cones, cylinders, and planes. A map projection systematically projects locations from the surface of a spheroid to representative positions on a flat surface using mathematical algorithms.

The first step in projecting from one surface to another is creating one or more points of contact. Each contact is called a point (or line) of tangency. A planar projection is tangential to the globe at one point. Tangential cones and cylinders touch the globe along a line. If the projection surface intersects the globe instead of merely touching its surface, the resulting projection is a secant rather than a tangent case. Whether the contact is tangent or secant, the contact points or lines are significant because they define locations of zero distortion. Lines of true scale include the central meridian and standard parallels and are sometimes called standard lines. In general, distortion increases with the distance from the point of contact.

Many common map projections are classified according to the projection surface used: conic, cylindrical, or planar.

Projection types illustrated

Each of the main projection types—conic, cylindrical, and planar—are illustrated below.

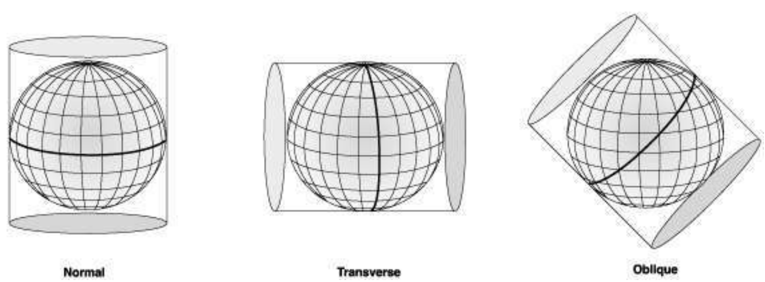
Conic (tangent)

A cone is placed over a globe. The cone and globe meet along a latitude line. This is the standard parallel. The cone is cut along the line of longitude that is opposite the central meridian and flattened into a plane.

Conic (secant)

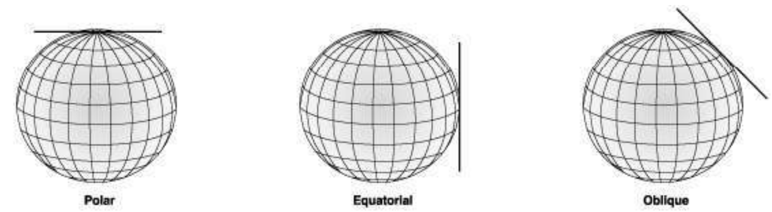
A cone is placed over a globe but cuts through the surface. The cone and globe meet along two latitude lines. These are the standard parallels. The cone is cut along the line of longitude that is opposite the central meridian and flattened into a plane.

Cylindrical aspects



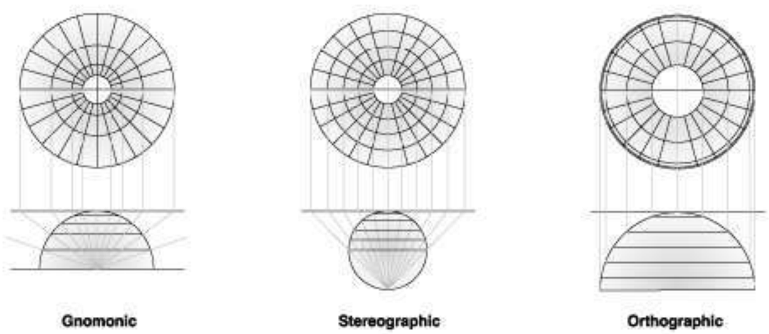
A cylinder is placed over a globe. The cylinder can touch the globe along a line of latitude (normal case), a line of longitude (transverse case), or another line (oblique case).

Planar aspects



A plane is placed over a globe. The plane can touch the globe at the pole (polar case), the equator (equatorial case), or another line (oblique case).

Polar aspect (different perspectives)



Azimuthal, or planar projections can have different perspective points. The gnomonic projection's point is at the center of the globe. The opposite side of the globe from the point of contact is used for a stereographic projection. The perspective point for an orthographic projection is at infinity.

Creation of a Map Projection

The creation of a map projection involves three steps in which information is lost in each step:

1. selection of a model for the shape of the earth or round body (choosing between a sphere or ellipsoid)
2. transform geographic coordinates (longitude and latitude) to plane coordinates (eastings and northings).
3. reduce the scale (in manual cartography this step came second, in digital cartography it comes last)

Metric properties of maps

Maps assume that the viewer has an orthogonal view of the map (they are looking straight down on every point). This is also called a perpendicular view or normal view. The metric properties or a map are

- area
- shape
- direction
- distance
- scale

Choosing a projection surface

If a surface can be transformed onto another surface without stretching, tearing, or shrinking, then the surface is said to be an applicable surface. The sphere or ellipsoid are not applicable with a plane surface so any projection that attempts to project them on a flat sheet will have to distort the image (similar to the impossibility of making a flat sheet from an orange peel). A surface that can be unfolded or unrolled into a flat plane or sheet without stretching, tearing or shrinking is called a 'developable surface'. The cylinder, cone and of course the plane are all developable surfaces since they can be unfolded into a flat sheet without distorting the projected image (although the original projection of the earth's surface on the cylinder or cone would be distorted).

Orientation of the projection

Once a choice is made between using a cylinder or cone is made, the orientation for that shape must be chosen (how the cylinder or cone is "placed" on the earth). The orientation of the projection surface can be normal (inline with the earth's axis), transverse (at right angles to the earth's axis) or oblique (any angle in between). These surfaces may also be either tangent or secant to the sphere or ellipsoid.

Datums

While a spheroid approximates the shape of the earth, a datum defines the position of the spheroid relative to the center of the earth. A datum provides a frame of reference for measuring locations on the surface of the earth. It defines the origin and orientation of latitude and longitude lines.

Whenever you change the datum, or more correctly, the geographic coordinate system, the coordinate values of your data will change. Here are the coordinates in degrees/minutes/seconds (DMS) of a control point in Redlands, California, on the North American Datum of 1983 (NAD 1983 or NAD83):

34 01 43.77884 -117 12 57.75961

Here's the same point on the North American Datum of 1927 (NAD 1927 or NAD27):

34 01 43.72995 -117 12 54.61539

The longitude value differs by approximately 3 seconds, while the latitude value differs by about 0.05 seconds.

NAD 1983 and the World Geodetic System of 1984 (WGS 1984) are identical for most applications. Here are the coordinates for the same control point based on WGS 1984:

34 01 43.778837 -117 12 57.75961

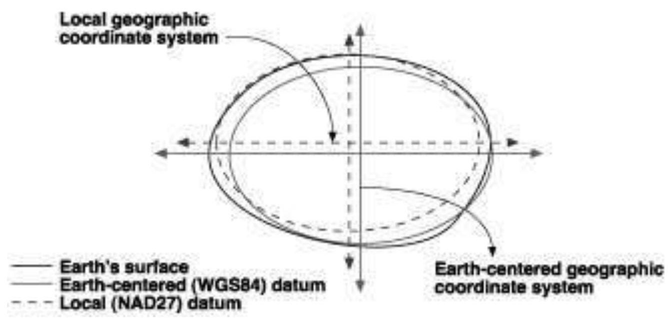
Geocentric datums

In the last 15 years, satellite data has provided geodesists with new measurements to define the best earth-fitting spheroid, which relates coordinates to the earth's center of mass. An earth-centered, or geocentric, datum uses the earth's center of mass as the origin. The most recently developed and widely used datum is WGS

1984. It serves as the framework for locational measurement worldwide.

Local datums

A local datum aligns its spheroid to closely fit the earth's surface in a particular area. A point on the surface of the spheroid is matched to a particular position on the surface of the earth. This point is known as the origin point of the datum. The coordinates of the origin point are fixed, and all other points are calculated from it.



The coordinate system origin of a local datum is not at the center of the earth. The center of the spheroid of a local datum is offset from the earth's center. NAD 1927 and the European Datum of 1950 (ED 1950) are local datums. NAD 1927 is designed to fit North America reasonably well, while ED 1950 was created for use in Europe. Because a local datum aligns its spheroid so closely to a particular area on the earth's surface, it's not suitable for use outside the area for which it was designed.

DATA TYPES AND DATA INPUTS: SPATIAL DATA AND ATTRIBUTES

For a GIS to be useful it must be capable of receiving and producing information in an effective manner.

The data input and output functions are the means by which a GIS communicates with the world outside.

The objective in defining GIS input and output requirements is to identify the mix of equipment and methods needed to meet the required level of performance and quality. No one device or approach is optimum for all situations.

DATA INPUT: The procedure of encoding data into a computer-readable form and writing the data to the GIS database.

Data entry is usually the major bottleneck in implementing a GIS. The initial cost of building the database is commonly 5 to 10 times to cost of the GIS hardware and software.

The creation of an *accurate* and *well-documented* database is critical to the operation of the GIS.

Accurate information can only be generated if the data on which it is based were accurate to begin with.

Data quality information includes the date of collection, the positional accuracy, completeness, and the method used to collect and encode the data. (Discussed in detail in Ch. 5)

There are two types of data to be entered into a GIS: Spatial data and the associated non-spatial attribute data.

The spatial data represents the geographical location of the features

The non-spatial attribute data provide descriptive information like the name of a street, salinity of the lake or the type of tree stand.

The non-spatial attribute data must be logically attached to the features they describe.

There are five types of data entry systems commonly used in a GIS:

- keyboard entry
- coordinate geometry
- manual digitizing
- scanning
- input of existing digital files

Keyboard entry: involves manually entering the data at a computer terminal. Attribute data are commonly input by keyboard whereas spatial data are rarely input this way.

Keyboard entry may also be used during manual digitizing to enter the attribute information. However this is usually more efficiently handled as a separate operation.

Roads files versus the census file — roads file will use codes for the various road types while the census file uses exact numbers for things like total population, age range, etc.

Coordinate Geometry (COGO): involves entering survey data using a keyboard. From these data the coordinate of spatial features are calculated. This produces a very high level of precision and accuracy which is needed in a cadastral system.

For a city with 100,000 parcels, it would cost approximately $1 - $1.50 per parcel or $100,000 to $150,000 to digitize the parcels manually. COGO procedures are commonly 6 times and can be up to 20 times more expensive than manual digitizing.

Surveyors and engineers want the higher accuracy of COGO for their applications. Planners and most others are happy with the lower accuracy provided by manual digitizing.

Manual Digitizing: The most widely used method for entering spatial data from maps. The map is mounted on a *digitizing tablet* and a hand held device termed a puck or cursor is used to trace each map feature. The position of the puck is accurately measured by the device to generate the coordinate data.

Digitizing surfaces range from 12 inches x 12 inches (digitizing tablet) to 36 x 48 (digitizing table) and on up.

The digitizing table electronically encodes the position of the pointing device with a precision of a fraction of a millimeter.

The most common digitizer uses a fine wire mesh grid embedded in the table. The cursor normally has 16 or more buttons that are used to operate the data entry and to enter attribute data.

The digitizing operation itself requires little computing power and so can be done without using the full GIS. A smaller, less expensive computer can be used to control the digitizing process and store the data. The data can later be transferred to the GIS for processing. The problem with this is having enough software for all the computers.

The efficiency of digitizing depends on the quality of the digitizing software and the skill of the operator. The process of tracing lines is time-consuming and error prone. The software can provide aids that substantially reduce the effort of detecting and correcting errors.

Attribute information may be entered during the digitizing process, but usually only as an identification number. The attribute information referenced to the same ID number is entered separately.

Manual digitizing is a tedious job. Operator fatigue (eye strain, back soreness, etc.) can seriously degrade the data quality. Managers must limit the number of hours an operator works at one time. A commonly used quality check is to produce a verification plot of the digitized data that is visually compared with the map from which the data were originally digitized.

Scanning: Scanning provides a faster means of data entry compared to manual digitizing.

In scanning, a digital image of the map is produced by moving an electronic detector across the surface of the map.

There are two types of scanner designs:

Flat-bed scanner: On a flat-bed scanner the map is placed on a flat scanning stage and the detectors move across the map in both the X and the Y directions (similar to copy machine).

Drum scanner: On a drum scanner, the map is mounted on a cylindrical drum which rotates while the detector moves horizontally across the map. The sensor motion provides movement in the X direction while the drum rotation provides movement in the Y direction.

The output from the scanner is a digital image. Usually the image is black and white but scanners can record color by scanning the same document three times using red, green and blue filters.

Inputting existing digital files: There are many companies and organizations on the market that provide or sell digital data files often in a format that can be read directly into a GIS. These digital data sets are priced at a fraction of the cost of digitizing existing maps.

Over the next decade, the increased availability of data should reduce the current high cost and lengthy production times needed to develop digital geographic data bases.

Scanning Versus Manual Digitizing

Scanning is being used by many organizations, yet the subject is very controversial. One reason for the questions on data accuracy is that rigorous trials are few and of necessity are specific to the organization and application.

Data entry using scanning is claimed to be 5 to 10 times (or more) faster than digitizing.

However maps normally must be redrafted before they can be scanned or the color separates must be scanned.

Redrafting is often considered to be a major disadvantage of the scanning option. Redrafting, although time consuming, does not necessarily add to the cost of the data conversion process. Redrafting can reduce the total cost of both scanning and manual digitizing. For example, studies by the US Forest Service have shown that a "map preparation" step before the manual digitizing is done can reduce the overall digital encoding costs by as much as 50%.

While a scanning system is for the most part automated, and requires less highly trained personnel, more complex equipment must be maintained, more sophisticated software must be written or purchased and there are most steps in the process.

Scanners are more expensive than digitizing tables. A 60 x 44 inch digitizing table can cost between $3000 and $8000. A high quality scanner will cost $100,000. The higher equipment costs can be justified if there is a great deal of production that needs to be done.

Most GIS software packages include a digitizing software capability, but separate special-purpose software is needed to operate a scanning system. Scanning works best with maps that are very clean, simple, and do not contain extraneous information. Scanning is most cost-effective for maps with large numbers of polygons (1000 or more) and maps with a large number of irregularly shaped features such as lines and odd polygons.

Manual digitizing tends to be more cost-effective when there are relatively few maps that are not in a form that can be scanned. Maps that require a lot of interpretation do not need to be scanned.

There is a strong demand for faster, more cost effective data entry methods. Hundreds of computer operators with thousands of maps are not the answer. Although scanning will never replace manual digitizing, as more and more scanners are used, the technology will become better and better.

Direct Use Of Raster Scanned Images

Much of the difficulty in using raster scanning to enter map information is the extraction of points, lines and polygons from the raster data.

In some cases, the raster image is only needed as a background on which to overlay other geographic information.

Air photos, satellite imagery, and scanned map images can be stored and presented in this way.

For example, if a raster satellite image is displayed on a screen, a vector map can be overlaid and then updated or a totally new map created by digitizing on the screen.

Using a raster image as a background can be an effective solution when a relatively small amount of data needs to be extracted but a large area must be displayed in order to find the data.

Existing Digital Data

In the US and Canada, low-cost digital geographic information is becoming more readily available.

Data sets are being produced by the national mapping agencies and agencies responsible for the census and other nationwide statistical data. In the US these agencies include but are not limited to the USGS, US Census Bureau, and the DMA. Natural resources information is being converted to digital form at both the federal and state or provincial levels.

Since digital data sets are produced to satisfy a wide range of users, the cost of the data, currency and accuracy vary. The accuracy with which boundaries are drawn, the date of the information, and the method compilation may be sufficiently different to create errors when different data layers or adjacent map sheets within a data layer are used together.

Problems such as there may occur in any digital data set and must be identified and taken into account.

Private companies are also beginning to provide off-the-shelf database products. Although there may be difficulties, the cost of existing data is usually a fraction of the cost of creating anew data set.

The availability of inexpensive data sets will make GIS technology economically more attractive and easier to implement.

In the US the cartographic community has made a considerable effort to coordinate and standardize the production and distribution of digital geographic data.

At the federal level, the Federal Interagency Coordinating Committee on Digital Cartography (FICCDC) was formed in 1983 for this purpose. Over 14 organizations participate in the Committee, which holds regular meetings and produces a newsletter and a variety of reports.

Now we are going to discuss examples of data sets available from these federal agencies.

Base Cartographic Data

Base cartographic data include the topographic and planimetric information usually portrayed on a map.

Topographic data are those data that portray relief, such as elevation contours and spot heights.

Planimetric data include roads and streams, as well as cultural data such as administrative and political boundaries, cities, and towns.

Often these data sets are digitized version of an existing map series with each type of information such as the elevation contours, assigned to separate data layers. Base cartographic data sets are produced in two formats: *Graphics and topologically-structured.*

Graphics format is essentially the line and point features digitized in vector format. In this form, the map can be easily updated or modified to produced special purpose maps.

These data sets are well suited for the CAD systems used in digital mapping. However, they are severely limited by the lack of topological structuring.

A commonly used interchange format is the SIF (Standard Interchange Format) developed by the digital mapping industry for transferring lines, points, curves, and symbols.

These data sets can be incorporated into a GIS but there can be a lot of problems associated with it. For example, the data files often have not been checked for topological consistency.

They may contain such inconsistencies as lines that do not met precisely, that overshoot or under shoot the correct connection point. The may be missing lines or gaps that create polygons that are not closed.

For use in a vector GIS these files must be clean and topologically structured.

Topologically-Structured Format is designed to encode geographic information in a form better suited for spatial analysis and other geographic studies. Most GISs are designed to use topologically structured data.

The USGS Digital Line Graph (DLG) data set is an example of topologically structured data. This cartographic data set has been developed from previous mapping efforts at the 1:2 million scale and more recently at the 1:100,000 and 1:24,000 scales.

The older 1:2 million data includes transportation , hydrography, and political boundary maps.

The 1:100,000 scale data sets for hydrography and transportation have been completed for the entire US while the political boundaries and Public Land Survey System are still being developed.

The 1:24,000 series will include the PLSS, political boundaries, transportation, hydrography, and contour data layers.

These data sets represent a comprehensive, standardized inexpensive and publicly available source of digital information.

The complete coverage (at the 1:100,000 scale) makes it possible to assemble large-area data bases quickly and at a low cost.

Land Use / Land Cover Data

The USGS has developed a LU/LC data set compiled from 1:58,000 color infrared aerial photography and mapped at the 1:250,000 scale.

The data sets were generated by both manual digitizing and scan digitizing.

The LU/LC classes include urban areas, agricultural land, rangeland, forest, wetlands, barren land and tundra.

Associated maps provide political boundaries, hydrological units (watershed boundaries), federal land ownership, and census subdivisions. Data are available for about 75% of the US. A separate file is being developed for Alaska using a different classification scheme and automated classification of digital satellite imagery.

Census-related Data Sets

In Canada and the US, the agencies responsible for disseminating census data provide a number of digital data sets that can be input to a GIS.

Census and other statistical data are provided in the form of attribute data sets coded by geographic location.

Enumeration districts, street addresses, postal codes, census tracts and other similar codes are used.

Spatial data sets are provided that can be linked to the attribute datasets by means of these area codes.

Street networks in metropolitan areas, census tract boundaries, and political boundaries are examples of the spatial data sets commonly available.

The spatial and attribute data sets are sued together to produce special purpose maps and to retrieve information for selection geographic areas. They are also used for more specialized analyses including *address matching, district delineation, and network analysis.*

Address Matching is the technique of linking data from separate files by means of a common attribute, the street address. For example, welfare case records may include the name and the address of each recipient but not the census tract. The census tract information can be retrieved from the spatial data file by using the address as a key to find the data in the other file.

District Delineation is a procedure that defines compact areas based on one ore more attributes. For example, it can be used to divide an area into electoral districts that each have about the same population. Conceptually, this involves starting at one point and enlarging the area until it encompasses the specified number of people, then a new district is started and the process is repeated.

The population information would be retrieved from the attribute data file and the information needed to define and enlarge the district boundaries would be retrieved from the spatial data file.

The district delineation procedure is used to define police and fire service districts, school districts, and commercial market areas.

Network Analysis is used to optimize transportation routing such as bus routes and emergency vehicle dispatching.

This procedure takes into account the length of each transportation segment and facts that affect the speed of travel or the quantity of material that can be carried. Sophisticated systems can take into account the effects of rush hour traffic, road closures, and vehicle availability in order to make the best assignment of delivery vehicles and routing.

GBF/DIME And Tiger Files

The US Census Bureau developed a geographic coding system to automate the processing of census questionnaires. This system, called GBF/DIME has been used since 1970.

The acronym stands form *Geographic Base File/Dual Independent Map Encoding system*. The files are topologically structured and were produced for 350 major cities and suburbs across the US. The spatial data included street networks, street addresses, political boundaries, and major hydrological features.

One of the benefits of this file was that census data could be easily aggregated by geographic regions for reporting purposes. Local governments found that the GBF/DIME files were inexpensive data sources for their GIS. Digital street maps could be produced from the data and after editing could be used as digital base maps for municipal applications.

However, the GBF/DIME files were not designed to be used as a digital map base and have some limitations. *Firs*t, the data do not accurately show the shape of the streets because each segment is a straight line connecting two intersections and therefore curved lines become straight lines.

Secondly, the address range is provided for each street segment but the geographic position of each address location is not included. In preparation for the 1990 Census, the Bureau of the Census developed the TIGER files (*Topologically Integrated Geographic Encoding and Referencing System*) to replace the GBF/DIME system. The TIGER overcame many of the limitations of the earlier system. It covers the 50 states, DC, Puerto Rico, the Virgin Islands of the US, and the outlying areas of the Pacific over which the US has jurisdiction.

Attribute data in the TIGER file include feature names, political and statistical geographic area codes (such as county, incorporated place, census tract and block number) and potential address ranges, and zip codes for that portion of the file. The Census Bureau no longer supports the DIME files. The TIGER files can be easily integrated into an existing GIS data base by file matching, using the geographic area codes as match keys.

Digital Elevation Data

Digital elevation data are a set of elevation measurements for locations distributed over the land surface. They are used to analyze the topography (surface features) of an area.

Various terms have been used to refer to digital elevation data and its derivatives:

- Digital Terrain Data Digital Terrain Models Digital Elevation Model
- Digital Terrain Elevation Data

Digital elevation data are used in a wide range of engineering, planning, and military applications. For example, the are used to:

- Calculate cut-and-fill operations for road construction;
- Calculate the area that would be flooded by a hydroelectric dam;
- Analyze and delineate area that can be seen from a location in the terrain;
- Intervisibility can also be used to plan route locations for roadways;
- Optimize the location of radar antennas or microwave towers; or

- Define the viewshed of an area.

The methods used to capture and store elevation data can be grouped into four basic approaches:

* A regular grid contours profiles
* Triangulated Irregular Network (TIN)
* Digital elevation data are generated from existing contour maps, by photogrammetric analysis of stereo aerial photographs, or more recently by automated analysis of stereo satellite data.
* DTM data are most commonly provided in grid format in which an elevation value is stored for each of a set of regularly spaced ground positions. Each data point represents the elevation of the grid cell in which it is located.

One of the limitations of the raster form of representation is that the same density of elevation points is used for the entire coverage area.

Ideally, the data points would be more closely spaced in complex terrain and sparsely distributed over more level areas. A number of methods have been developed to provide a variable point density.

One method is to use a variable grid cell spacing to accommodate a variable density of points, with smaller cell sizes being used to capture the detail in more complex terrain. *A second approach* has been to use irregularly spaced elevation points and represent the topography by a network of triangular facets. In this way, elevation data can be stored and manipulated using a vector representation. The TIN is produced from a set of irregularly spaced elevation points. A network of triangular facets is fit to these points. The coordinate positions and elevations of the three points forming the vertices of each triangular facet are used to calculate such terrain parameters as the slope and aspect.

The advantage of a TIN compared with a gridded representation is that the TIN can use fewer points, capture the critical points that define discontinuities like ridge crests, and can be topologically encoded so that adjacency analyses are more easily done.

A third way to digitally represent a topographic surface is by development of a profile showing the elevation of points along a series of parallel lines.

Elevation values should be recorded at all breaks in slope and at scattered points in level terrain. If the profiles are constructed from a topographic map, the elevation values can only be taken where the profile crosses a contour line.

The fourth approach is to digitize contour lines. Here the topographic surface is represented by series of elevation points taken along the individual contours. Although elevation data can be converted from one format to another, each time the data are converted some information is lost reducing the detail to the topographic surface.

Digital elevation data is available in the US and was first produced by the Defense Mapping Agency. They were produced by scanning the contour overlays for 1:250,000 scale topographic maps.

These data have an accuracy of 15 m in level terrain, 30m in moderate terrain, and 60 m in steep terrain.

The data are sold by the map sheet as 1 degree x 1 degree blocks and are available for the entire US.

The USGS plans to progressively upgrade the accuracy of this data set and is also producing a higher accuracy DTM file with a 30m sampling interval. The data are maintained in two datasets; one with a +7m accuracy and the other with a +7 - +15m accuracy. These data are available for about 30% of the US and are sold by 7.5 minute quad sheets.

The unit price for these data decrease with the number of DTs purchased. Prices for orders of six or more DTM consist of a base charge of $90 and $7 for each additional unit.

Data Output

Output is the procedure by which information from the GIS is presented in a form suitable to the user. Data are output in one of three formats: Hardcopy, Softcopy and electronic.

Hardcopy outputs are permanent means of display. The information is printed on paper, mylar, photographic film or other similar materials.

Softcopy output is in the format viewed on a computer monitor. Softcopy outputs are used to allow operator interaction and to preview data before final output. A Softcopy output can be changed interactively but he view is restricted by the size of the monitor.

The hardcopy output takes longer to produce and requires more expensive equipment. However, it is a permanent record.

Output in electronic formats consists of computer-compatible files.

Data Input Techniques

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.

Consensus within the GIS community indicates that scanners work best when the information on a map is kept very clean, very simple, and uncluttered with graphic symbology.

The sheer cost of scanning usually eliminates the possibility of using scanning methods for data capture in most GIS implementations. Large data capture shops and government agencies are those most likely to be using scanning technology.

Currently, general consensus is that the quality of data captured from scanning devices is not substantial enough to justify the cost of using scanning technology. However, major breakthroughs are being made in the field, with scanning techniques and with capabilities to automatically clean and prepare scanned data for topological encoding. These include a variety of *line following* and *text recognition* techniques. Users should be aware that this technology has great potential in the years to come, particularly for larger GIS installations.

Coordinate Geometry

A third technique for the input of spatial data involves the calculation and entry of coordinates using coordinate geometry (COGO) procedures. This involves entering, from survey data, the explicit measurement of features from some known monument. This input technique is obviously very costly and labour intensive. In fact, it is rarely used for natural resource applications in GIS. This method is useful for creating very precise cartographic definitions of property, and accordingly is more appropriate for land records management at the cadastral or municipal scale.

Conversion of Existing Digital Data

A fourth technique that is becoming increasingly popular for data input is the conversion of existing digital data. A variety of spatial data, including digital maps, are openly available from a wide range of government and private sources. The most common digital data to be used in a GIS is data from CAD systems. A number of data conversion programs exist, mostly from GIS software vendors, to transform data from CAD formats to a raster or topological GIS data format. Several ad hoc standards for data exchange have been established in the market place. These are supplemented by a number of government distribution formats that have been developed. Given the wide variety of data formats that exist, most GIS vendors have developed and provide data exchange/conversion software to go from their format to those considered common in the market place.

Most GIS software vendors also provide an ASCII data exchange format specific to their product, and a programming subroutine library that will allow users to write their own data conversion routines to fulfil their own specific needs. As digital data becomes more readily available this capability becomes a necessity for any GIS. Data conversion from existing digital data is not a problem for most technical persons in the GIS field. However, for smaller GIS installations who have limited access to a *GIS analyst* this can be a major stumbling block in getting a GIS operational. Government agencies are usually a good source for technical information on data conversion requirements.

GIS Data: Spatial vs Attributes

GIS Data is the key component of a GIS and has two general types: Spatial and Attribute data.

Spatial data are used to provide the visual representation of a geographic space and is stored as raster and vector types. Hence, this data is a combination of location data and a value data to render a map, for example.

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. The nominal level is the lowest level of measurement for distinguishing features quantitatively using type or class (e.g. tree species). Ordinal data are ranked into hierarchies but does not show any magnitude of difference (e.g. city hierarchy). The interval measurement indicates the distance between the ranks of measured elements, but a starting point is arbitrarily assigned (e.g. Celsius Temperature). Ratio measurements, the highest level of measurements, includes an absolute starting point. Data of this category include property value and distance.

Attribute data is the detailed data used in combination with spatial data to create a GIS. The more available and appropriate attribute data used with spatial data, the more complete a GIS is as a management reporting and analysis tool.

Sources of Spatial & Attribute Data

Spatial data can be obtained from satellite images or scanned maps and similar resources. This data can then be digitised into vector data or maintained as raster graphic data. Essentially, any format of a geographical image with location or co-ordinate points can be used as spatial data.

Attribute data can be obtained from a number of sources or data can be captured specifically for you application. Some popular sources of attribute data are from town planning and management departments, policing and fire departments, environmental groups, online media.

DATA INPUT, SCANNING, DIGITIZATION, ERROR CORRECTIONS AND TOPOLOGY

GIS and scanning technology

Maps are generally considered the backbone of any GIS activity. But many a time paper maps are not easily available in a form that can be readily used by the computers. Most of the paper maps had been prepared on the basis of old conventional surveys. New maps can be produced using improved technologies but this requires time as it increases the volume of work. Thus, we have to resort to the available maps. These paper maps have to be first converted into a digital format usable by the computer. This is a critical step as the quality of the analog document must be preserved in the transition to the computer domain. The technology used for this kind of conversions is known as scanning and the instrument used for this kind of operation is known as a scanner. A scanner can be thought of as an electronic input device that converts analog information of a document like a map, photograph or an overlay into a digital format that can be used by the computer.

Scanning automatically captures map features, text, and symbols as individual cells, or pixels, and produces an automated image.

Working of a Scanner

The most important component inside a scanner is the scanner head which can move along the length of the scanner. The scanner head contains either a charged-couple device (CCD) sensor or a contact image (CIS) sensor. A CCD consists of a number of photosensitive cells or pixels packed together on a chip. The most advanced large format scanners use CCD's with 8000 pixels per chip for providing a very good image quality.

While scanning a bright white light from the scanner strikes the image to be scanned and is reflected onto the photosensitive surface of the sensor placed on the scanner head. Each pixel transfers a graytone value (values given to the different shades of black in the image ranging from 0 (black) – 255 (white) i.e. 256 values to the scanboard (software).

The software interprets the value in terms of 0 (Black) or 1 (white), thereby, forming a monochrome image of the scanned portion. As the head moves ahead, it scans the image in tiny strips and the sensor continues to store the information in a sequential fashion.

The software running the scanner pierces together the information from the sensor into a digital form of the image. This type of scanning is known as one pass scanning.

Scanning a colour image is slightly different in which the scanner head has to scan the same image for three different colours i.e. red, green, blue. In older colour scanners, this was accomplished by scanning the same area three times over for the three different colours. This type of scanner is known as three-pass scanner. However, most of the colour scanners now scan in one pass scanning all the three colours in one go by using colour filters. In principle, a colour CCD works in the same way as a monochrome CCD. But in this each colour is constructed by mixing red, green and blue. Thus, a 24-bit RGB CCD presents each pixel by 24 bits of information. Usually, a scanner using these three colours (in full 24 RGB mode) can create up to 16.8 million colours.

Nowadays a new technology: full width, single-line contact sensor array scanning has emerged in which the document to be scanned passes under a line of LED's which capture the image. This new technology enables the scanner to operate at previously unattainable speeds.

Types of Scanners

There are several different types of scanners performing the same job but handling the job differently using different technologies and producing results depending on their varying capabilities.

Hand-held scanners although portable, can only scan images up to about four inches wide. They require a very steady hand for moving the scan head over the document. They are useful for scanning small logos or signatures and are virtually of no use for scanning maps and photographs.

Hand held Scanner

The most commonly used scanner is a flatbed scanner also known as desktop scanner. It has a glass plate on which the picture or the document is placed. The scanner head placed beneath the glass plate moves across the picture and the result is a good quality scanned image. For scanning large maps or toposheets wide format flatbed scanners can be used.

Flatbed Scanner

Then there are the drum scanners which are mostly used by the printing professionals. In this type of scanner, the image or the document is placed on a glass cylinder that rotates at very high speeds around a centrally located sensor containing photo-multiplier tube instead of a CCD to scan. Prior to the advances in the field of sheet fed scanners, the drum scanners were extensively used for scanning maps and other documents.

Drum Scanner

Finally, there are the Sheet fed scanners which work on a principle similar to that of a fax machine. In this, the document to be scanned is moved past the scanning head and the digital form of the image is obtained. The disadvantage of this type of scanner is that it can only scan loose sheets and the scanned image can easily become distorted if the document is not handled properly while scanning. However, the new generation of the wide format sheet fed scanners has overcome this problem and have become indispensable for scanning maps, imageries and other large sized documents.

Digitizing Errors in GIS

Digitizing in GIS is the process of converting geographic data either from a hardcopy or a scanned image into vector data by tracing the features. During the digitzing process, features from the traced map or image are captured as coordinates in either point, line, or polygon format.

Types of Digitizing in GIS

There are several types of digitizing methods. Manual digitizing involves tracing geographic features from an external digitizing tablet using a puck (a type of mouse specialized for tracing and capturing geographic features from the tablet). Heads up digitizing (also referred to as on-screen digitizing) is the method of tracing geographic features from another dataset (usually an aerial, satellite image, or scanned image of a map) directly on the computer screen. Automated digitizing involves using image processing software that contains pattern recognition technology to generated vectors.

Types of Digitizing Errors in GIS

Since most common methods of digitizing involve the interpretation of geographic features via the human hand, there are several types of errors that can occur during the course of capturing the data. The type of error that occurs when the feature is not captured properly is called a positional error, as opposed to attribute errors where information about the feature capture is inaccurate or false. These positional error types are outlined below, and a visualization of the different methods is shown at the bottom of this chapter.

During the digitizing process, vectors are connected to other lines by a node, which marks the point of intersection. Vertices are defining points along the shape of an unbroken line. All lines have a starting point known as a starting node and an ending node. If the line is not a straight line, then any bends and curves on that line are defined by vertices (vertex for a singular bend). Any intersection of two lines is denoted by node at the point of the intersection.

Topology

Topology expresses explicitly the spatial relationships between connecting or adjacent vector features (points, polylines and polygons) in a GIS, such as two lines meeting perfectly at a point and directed line having an explicit left and right side.

Topological or topology based data are useful for detecting and correcting digitizing error in geographic data set and are necessary for some GIS analyses.

Topologic data structures help insure that information is not unnecessarily repeated. The database stores one line only in order to represent a boundary (as opposed to two lines, one for each polygon). The database tells us that the line is the "left side" of one polygon and the "right side" of the adjacent polygon.

Topology is the study of those properties of geometric objects that remain invariant under certain transformations such as bending or stretching.

Topology is often explained through graph theory. Topology has least two main advantages.

(i) The assurance of data quality

(ii) Enhance GIS analysis

Topological relationships are built from simple elements into complex elements: points (simplest elements), arcs (sets of connected points), areas (sets of connected arcs), and routes (sets of sections, which are arcs or portions of arcs).

Components of Topology

Topology has three basic components:

1. Connectivity (Arc – Node Topology):
   - Points along an arc that define its shape are called Vertices.
   - Endpoints of the arc are called Nodes.
   - Arcs join only at the Nodes.
2. Area Definition / Containment (Polygon – Arc Topology):
   - An enclosed polygon has a measurable area.
   - Lists of arcs define boundaries and closed areas are maintained.
   - Polygons are represented as a series of (x, y) coordinates that connect to define an area.
3. Contiguity:
   - Every arc has a direction
   - A GIS maintains a list of Polygons on the left and right side of each arc.
   - The computer then uses this information to determine which features are next to one another.

2

Data Models

INTRODUCTION

A data model is used to organize data. A data model captures the cardinality and referential integrity rules needed to ensure that the data is of good quality for the users. A data model has 3 uses in an application which are getting data in, integrating data and getting data out. A data model is also used as a communication tool for teams to communicate within the team on how data is organized and between teams.

Overview

Managing large quantities of structured and unstructured data is a primary function of information systems. Data models describe structured data for storage in data management systems such as relational databases. They typically do not describe unstructured data, such as word processing documents, email messages, pictures, digital audio, and video.

The Role of Data Models

The main aim of data models is to support the development of information systems by providing the definition and format of data. According to West and Fowler (1999) "if this is done consistently across systems then compatibility of data can be achieved. If the same data structures are used to store and access data then different applications can share data. However, systems and interfaces often cost more than they should, to build, operate, and maintain. They may also constrain the business rather than support it. A major cause is that the quality of the data models implemented in systems and interfaces is poor".

- "Business rules, specific to how things are done in a particular place, are often fixed in the structure of a data model. This means that small changes in the way business is conducted lead to large changes in computer systems and interfaces".
- "Entity types are often not identified, or incorrectly identified. This can lead to replication of data, data structure, and functionality, together with the attendant costs of that duplication in development and maintenance".
- "Data models for different systems are arbitrarily different. The result of this is that complex interfaces are required between systems that share data. These interfaces can account for between 25-70% of the cost of current systems".
- "Data cannot be shared electronically with customers and suppliers, because the structure and meaning of data has not been standardised. For example, engineering design data and drawings for process plant are still sometimes exchanged on paper".

The reason for these problems is a lack of standards that will ensure that data models will both meet business needs and be consistent.

Three Perspectives

A data model *instance* may be one of three kinds according to ANSI in 1975:

- Conceptual schema : describes the semantics of a domain, being the scope of the model. For example, it may be a model of the interest area of an organization or industry. This consists of entity classes, representing kinds of things of significance in the domain, and relationships assertions about associations between pairs of entity classes. A conceptual schema specifies the kinds of facts or propositions that can be expressed using the model. In that sense, it defines the allowed expressions in an artificial 'language' with a scope that is limited by the scope of the model. The use of conceptual schema has evolved to become a powerful communication tool with business users. Often called a subject area model (SAM) or high-level data model (HDM), this model is used to communicate core data concepts, rules, and definitions to a business user as part of an overall application development or enterprise initiative. The number of objects should be very small and focused on key concepts. Try to limit this model to one page, although for extremely large organizations or complex projects, the model might span two or more pages.

- Logical schema : describes the semantics, as represented by a particular data manipulation technology. This consists of descriptions of tables and columns, object oriented classes, and XML tags, among other things.
- Physical schema : describes the physical means by which data are stored. This is concerned with partitions, CPUs, tablespaces, and the like.

The significance of this approach, according to ANSI, is that it allows the three perspectives to be relatively independent of each other. Storage technology can change without affecting either the logical or the conceptual model.

The table/column structure can change without (necessarily) affecting the conceptual model. In each case, of course, the structures must remain consistent with the other model.

The table/column structure may be different from a direct translation of the entity classes and attributes, but it must ultimately carry out the objectives of the conceptual entity class structure. Early phases of many software development projects emphasize the design of a conceptual data model. Such a design can be detailed into a logical data model. In later stages, this model may be translated into physical data model. However, it is also possible to implement a conceptual model directly.

History

One of the earliest pioneering works in modelling information systems has been done by Young and Kent (1958), who argued for "a precise and abstract way of specifying the informational and time characteristics of a data processing problem". They wanted to create "a notation that should enable the analyst to organize the problem around any piece of hardware".

Their work was a first effort to create an abstract specification and invariant basis for designing different alternative implementations using different hardware components. A next step in IS modelling was taken by CODASYL, an IT industry consortium formed in 1959, who essentially aimed at the same thing as Young and Kent: the development of "a proper structure for machine independent problem definition language, at the system level of data processing".

This led to the development of a specific IS information algebra. In the 1960s data modelling gained more significance with the initiation of the management information system (MIS) concept.

According to Leondes (2002), "during that time, the information system provided the data and information for management purposes. The first generation database system, called Integrated Data Store (IDS), was designed by Charles Bachman at General Electric. Two famous database models, the network data model and the hierarchical data model, were proposed during this period of time". Towards the end of the 1960s Edgar F. Codd worked out his theories of data arrangement, and proposed the relational model for database management based on first-order predicate logic.

In the 1970s entity relationship modelling emerged as a new type of conceptual data modelling, originally proposed in 1976 by Peter Chen.

Entity relationship models were being used in the first stage of information system design during the requirements analysis to describe information needs or the type of information that is to be stored in a database. This technique can describe any ontology, i.e., an overview and classification of concepts and their relationships, for a certain area of interest. In the 1970s G.M. Nijssen developed "Natural Language Information Analysis Method" (NIAM) method, and developed this in the 1980s in cooperation with Terry Halpin into Object-Role Modelling (ORM).

Further in the 1980s according to Jan L. Harrington (2000) "the development of the object-oriented paradigm brought about a fundamental change in the way we look at data and the procedures that operate on data.

Traditionally, data and procedures have been stored separately: the data and their relationship in a database, the procedures in an application program. Object orientation, however, combined an entity's procedure with its data."

VECTOR AND RASTER DATA MODELS

Vector data

Vector data is split into three types: polygon, line (or arc) and point data. Polygons are used to represent areas such as the boundary of a city (on a large scale map), lake, or forest. Polygon features are two dimensional and therefore can be used to measure the area and perimeter of a geographic feature. Polygon features are most commonly distinguished using either a thematic mapping symbology (color schemes), patterns, or in the case of numeric gradation, a color gradation scheme could be used.

Line (or arc) data is used to represent linear features. Common examples would be rivers, trails, and streets. Line features only have one dimension and therefore can only be used to measure length. Line features have a starting and ending point. Common examples would be road centerlines and hydrology. Symbology most commonly used to distinguish arc features from one another are line types (solid lines versus dashed lines) and combinations using colors and line thicknesses. In the example below roads are distinguished from the stream network by designating the roads as a solid black line and the hydrology a dashed blue line.

Point data is most commonly used to represent nonadjacent features and to represent discrete data points. Points have zero dimensions, therefore you can measure neither length or area with this dataset. Examples would be schools, points of interest, and in the example below, bridge and culvert locations. Point features are also used to represent abstract points. For instance, point locations could represent city locations or place names.

Both line and point feature data represent polygon data at a much smaller scale. They help reduce clutter by simplifying data locations. As the features are zoomed in, the point location of a school is more realistically represented by a series of building footprints showing the physical location of the campus. Line features of a street centerline file only represent the physical location of the street. If a higher degree of spatial resolution is needed, a street curbwidth file would be used to show the width of the road as well as any features such as medians and right-of-ways (or sidewalks).

Raster Data

Raster data (also known as grid data) represents the fourth type of feature: surfaces. Raster data is cell-based and this data category also includes aerial and satellite imagery. There are two types of raster data: continuous and discrete. An example of discrete raster data is population density. Continuous data examples are temperature and elevation measurements. There are also three types of raster datasets: thematic data, spectral data, and pictures (imagery).

This example of a thematic raster dataset is called a Digital Elevation Model (DEM). Each cell presents a 30m pixel size with an elevation value assigned to that cell. The area shown is the Topanga Watershed in California and gives the viewer and understand of the topography of the region. Each cell contains one value representing the dominate value of that cell. Raster datasets are intrinsic to most spatial analysis. Data analysis such as extracting slope and aspect from Digital Elevation Models occurs with raster datasets. Spatial hydrology modeling such as extracting watersheds and flow lines also uses a raster-based system. Spectral data presents aerial or satellite imagery which is then often used to derive vegetation geologic information by classifying the spectral signatures of each type of feature.

What results from the effect of converting spatial data location information into a cell based raster format is called stairstepping. The name derives from the image of exactly that, the square cells along the borders of different value types look like a staircase viewed from the side.

Unlike vector data, raster data is formed by each cell receiving the value of the feature that dominates the cell. The stairstepping look comes from the transition of the cells from one value to another. In the image above the dark green cell represents chamise vegetation. This means that the dominate feature in that cell area was chamise vegetation. Other features such as developed land, water or other vegetation types may be present on the ground in that area. As the feature in the cell becomes more dominantly urban, the cell is attributed the value for developed land, hence the pink shading.

GIS Models

Geo-science specialists are busy with the modelling of real phenomena. A universal model to comprise all the aspects of reality is not practically realisable due to the high complexity of the real world. Different disciplines emphasise different aspects and only these aspects are included in the model. Thus a model considered good for the description of particular phenomena might be hardly appropriate for anothers. Different aspects and characteristics of real objects have led to the existence of several variations in object definition.

The term *GIS model* is utilised here to denote a data model of the real world. Being a data model for representing real-world objects, the GIS models have the components defined above, i.e. object types, relationships and attributes with corresponding generation rules, constraints and operations defined with them. A GIS model with the corresponding user interface constitutes a 3D GIS. Often the GIS content is specified as data about geometric (shape, size, location) and semantic characteristics (called attributes), spatial relationships and time. The *functionality* of GIS is then said to be the possibility to perform operations on data in order to analyse them.

The following sections present the necessary fundamental concepts related to the representation of objects, attributes and relationships in a GIS context.

Types of Objects

Since the interest in geo-sciences has traditionally been directed toward real objects with spatial extent, the differentiation between *spatial* and *non-spatial* objects is widely accepted. A spatial object stands for a real object having *geometric* and *thematic (semantic)* characteristics. Tempfli 1998c draws attention to a third group, i.e. *radiometric* characteristics of spatial objects. Current GIS models maintain only spatial objects. Non-spatial objects (organisations, departments, people, goods, etc.) either are the concern of Database Management Systems (DBMS) or are integrated in GIS as semantic characteristics of spatial objects. For example, the person, who owns a building (spatial) exists in a GIS as a semantic characteristics "owner" of the building. As will be discussed in this thesis, dealing with only non-spatial objects may be a drawback for some applications.

A further distinction is made between real objects with respect to distinctness of properties, i.e. objects with well defined spatial extent and properties and objects with unknown or non-well defined spatial extent and properties, i.e. objects with *discernible (determined)* and *indiscernible (undetermined)* boundaries. Usually, quite independent research is conducted in both areas. This may cause problems in the case of analysis where objects from the two groups are needed. Pilouk 1996 discusses the subject in details and advocates an integrated approach while modelling spatial objects.

Many real objects need the monitoring of some of their characteristics with respect to time. For example, growth of population may cause the fast conversion of land from rural to urban use, which can have an impact on the entire urban development. The geo-science specialist might need to store and analyse the changes in order to be able to predict and control the process. Similar problems have created the branch of temporal definitions of an object, according to which objects are subdivided into *spatio-temporal* and *non-spatio-temporal*. Raza et al 1998 present spatio-temporal-attribute objects (STAO) with three fundamental components, i.e. *location* (*spatial*), *attribute* (*aspatial*) and *time* (*temporal*). Cheng 1999 defines a real object by its three aspects *geometry*, *theme* and *time*.

More general definitions of real objects, regardless of the nature of the objects can be found in the object-oriented literature. Coad and Yourdon 1991 denote as an object every item of interest (real object, feature, process) and present a framework to investigate and classify the characteristics of the different objects.

Attributes

Attributes stand for the characteristics of real objects, which are important for the GIS model. The term attribute is frequently used to denote only semantic characteristics of objects. This thesis considers a wider meaning for attributes, i.e. they compile semantic, geometric and radiometric characteristics of objects. The geometric characteristics refer to position (and orientation), shape and size of real objects. The radiometric characteristics regard material, colour or reflectance. The semantic characteristics specify status, functionality, "meaning", usage, etc., of the real objects. This thesis concentrates on geometric characteristics and therefore some basic principles for representing them will be mentioned.

The description of geometric characteristics requires a priori clarification of the abstraction of space, the dimension of space and objects, and the method for representation. Conventionally, two approaches to space abstraction are utilised in modelling processes, i.e. field-oriented and object-oriented. The field-oriented approach assumes complete subdivision of the space into smaller, regular partitions, e.g. pixel. In the object-oriented, the space is "empty" and all the objects are places (embedded) in it, i.e. a lake in a 2D map. Both approaches have advantages and disadvantages and are appropriate for different applications. While the field-oriented approach better suits the representation of continuous phenomena, e.g. height fields, rainfalls, the object-oriented approach represents better discrete phenomena, e.g. buildings, roads.

The dimension of space is defined in mathematics as the number n of a sequence of n -real numbers (a_1, a_2,…,a_n), called an ordered n-tuple. All the ordered tuples are called n-space and are denoted by R^n. The basic idea is using a tuple of points to define a position in the space, e.g. a pair of points (x and y co-ordinates in Euclidean space) define a position in a plane, i.e. 2D space. The real world is usually represented as a 2D or 3D space. Within the space, the objects have their own dimensionality, i.e. they can be represented by one of the following generic types: 0D (points), 1D (lines), 2D (surfaces) and 3D (bodies) objects. The decision as to the representation of real objects (i.e. point, line, surface or body), is highly influenced by the purpose of the model.

The possible geometric representations are split into three large groups: i.e. raster, vector and constructive primitives. The "building blocks"(named constructive objects in this thesis) in the raster method are regular cells, e.g. pixels (in 2D space), voxels (in 3D space), which fill in the entire object. The representation is simple and easy to maintain, but produces a lot of data for storage, and the overall precision is low. The vector method is based on irregular n-cells, where n={1,2,3} composed of points with co-ordinates. In contrast to the raster method, the cells represent the boundaries of the objects. The last representation, known as Constructive Solid Geometry (CSG), uses irregular 3-cells. An object is represented by a CSG-tree, which holds information about the CSG primitives and the operations "gluing" them (i.e. Boolean operations). The approach is widely implemented in the manufacturing industry. In general, the irregular cells describing the boundaries of the objects allow more precise descriptions of shapes and spatial relationships compared to the raster cells and CSG primitives. Moreover, the visualisation algorithms operate with the boundaries of objects, i.e. CSG and raster representations have to be further processed to determine the boundary. Therefore many applications give preference to the vector approach of description. This thesis is based on a vector representation too.

The vector representations are often referred to as boundary representations (B-rep) or surface representations. A large number of spatial models are developed and implemented in GIS, CG and CAD systems based on irregular multidimensional cells. The names and construction rules of the cells in the different models usually vary. The simplest set of cells is the set of simplexes, i.e. 0-simplex (node), 1-simplex (arc), 2-simplex (triangle) and 3-simplex (tetrahedron). Composing simplexes, one can obtain more complex objects, i.e. simplicial complexes. Most of the models allow 1,2,3-cells with an arbitrary shape that imposes some supplementary constraints, e.g. planarity of faces, convex shape. The names vertex (point), edge, face (polygon) and solid (polyhedron) are then used in the literature to denote 0,1,2,3-cell.

Spatial Relationships

The particularity of GIS compared with other information systems is the maintenance of spatial relationships, i.e. the connections or interrelations between real objects in the geometric domain. The aspects of the spatial relationships are currently under investigations, i.e. approaches to representation, naming and equivalence.

Three different approaches to encoding spatial relationships are discussed in the literature, i.e. *metric*, *topology* and *order*. The metric is a pure computational approach, based on the comparison of numerical values related to the location of the objects in the space (i.e. the Euclidean space). For example, the spatial relationship between a house and a parcel (e.g. inside, outside, to the south) can be clarified by a metric operation point-in-polygon performed for each point constituting the footprint of the building. Since the metric is built on the notion of the *distance function*, which is dependent on the internal (finite) representation of numbers in the computers, the approach is computational expensive.

The order establishes a preference based on the mathematical relation “<” (*strict order*) or “£ ” (*partial order*), which allows an organisation of objects similar to a tree. For example, if a building is inside a parcel, the spatial relationship is represented as “building < parcel”. The applicability to representing spatial relationships is investigated by Kainz 1989 who argues that it has advantages in expressions of *inside/outside* relationships.

Topology allows the encoding of spatial relationships based on the *neighbourhoods* of objects regardless of the distance between them. The main property of topology, i.e. the invariance under topological transformations (i.e. rotation, scaling and translation) makes it appropriate for computer maintenance of spatial relationships. The following section discusses some basic topological issues that the thesis utilises.

Detection of Spatial Relationships

The development of a mathematical theory to categorise relations among spatial objects has been identified as an essential task to overcome the diversity and incompleteness of spatial relationships realised in different information systems. The intensive research in this area has led to the development of a framework based upon set theory and general topology principles and notions. The framework utilises the fundamental notions of general topology for topological primitives to investigate the interactions of the spatial objects. The topological primitives of a spatial object can be defined for each spatial model and hence the framework can be applied to any spatial model. The basic criterion to distinguish between different relations is the detection of empty and non-empty intersections between topological primitives. Depending on the number of the topological primitives considered, two *intersection models* were presented in the literature. The first idea is to investigate the intersection of interiors and boundaries of two objects. This results in 2^4=16 relations between two objects. Eight relations are given names, i.e. *disjoint*, *meet*, *contains*, *covers*, *inside*, *coveredBy*, *equal* and *overlap*).

For example, if the boundaries of the two objects intersect but the interiors do not, then the conclusion is that the objects *meet*. Apparently, many relations cannot be distinguished on the basis of only two topological primitives, therefore the evaluation of the exterior is adopted.

The number of detectable relations between two objects increases to 2^9=256. Despite the criticism (i.e. not all the relations are possible in reality, the intersections are not further investigated, many object intersections are topologically equivalent), the framework provides a systematic, easy-to-implement way of detecting spatial relations and is therefore used in the thesis. A slightly different approach is followed by Clementi et al 1993. Again, the three topological primitives are used but first the type of intersection is clarified and then a detailed evaluation of all the cells composing an object is performed. The approach claims a detection of a larger number of relations, however, at a high computational price.

Operations

Apart from the generic operations, GISs perform a number of specialised operations, some of which fall into the group of supporting operations, i.e. selection, navigation and specialisation. Plenty of classifications of required operations are presented in the literature that often confuses the reader. Usually, the classifications try to address three components of GIS, i.e. the geometric and semantic properties, and spatial relationships.

Goodchild 1987 specifies the spatial operations, which have to be performed in six groups:

- operations requiring access to semantic properties of one type of object, e.g. houses that belong to one owner
- operations requiring access to both semantic and location information, e.g. houses higher than three floors located to the north of the station
- operations which create object-pairs from one or more types of objects, e.g. all the buildings included in a given parcel
- operations which analyse semantics of object-pairs, e.g. the buildings in a given parcel that have one owner
- operations which require access to semantic and location information for more than one type of objects or object-pairs: spatial interaction modelling requires access to origin and destination of objects, as well as, their attributes.
- operations which create a new type of object from existing objects, e.g. Thiessen polygons or buffering polygons around lines.

Aronoff, 1995 summarises the required operations of a GIS as:

- retrieval operations, e.g. what is the current information about a particular building
- query operations, i.e. retrieval under a condition e.g. the houses cheaper than $50 000
- modelling operations, e.g. what will be the state of the data after ten years or the next year (predicting new information).

Further elaboration on needed retrieval and query operations results in subdivision into four groups:

- retrieval and query of the spatial data, i.e. transformation (geometric, format, geographical, conflation), and editing (edge matching)

- retrieval and query of the semantic data
- integrated analysis of spatial and semantic data, i.e. retrieval, classification, measurement, overlay operations
- neighbourhood operations (search, topographic operations, Thiessen polygons, interpolation, contour generation)
- connectivity operations (contiguity measurements, buffering, network, spread, seek, inter-visibility, illumination, perspective view)
- output formatting (map annotation, text labels, texture pattern and line style, graphic symbols).

In this thesis, we assume that GIS has to be able to perform selection, network and specialisation operations over the components it contains (e.g. geometric and semantic characteristics, and relationships). The result of these operations can be further processed to obtain a more specific result. Theoretically, the original operation and the further processing can be encapsulated in a new operation. Thus, we will define the following groups of GIS operations that are used in the thesis as follows:

- metric operations, i.e. distance, volume, area and length, are selection operations based on shape and size of objects and further processing
- position operations are selection operations based on location (no further processing)
- proximity operations are selection operations based on geometric characteristics and the creation of new object (i.e. buffer)
- relationship operations are selection operations based on spatial relationships (no further processing)
- visibility operations are a selection based on geometric characteristics and further processing
- semantic operations are selections based on semantic characteristics
- mixed operations, i.e. selections on the basis of geometric and semantic characteristics.

Molenaar 1998 specifies the *query* as a selection operation with three components: data type specification, conditions and operations that have to be performed on the data. The selection then can be performed on semantics, geometry or topology. For example, "select the buildings (data type) higher than 15m (condition) and show their ID (operation)". If the operations on data are not very complex, the definition of a query operation is similar to that given by Aronoff and Goodchild. Sophisticated operations on data may diminish the boundary between query and analysis. Therefore, a selection of data under conditions related to geometric, semantic characteristics or relationships, which are stored in the database, will be referred to here as *ad hoc query* or *query*. A selection of data that needs further processing is called *embedded query.* Embedded queries may be provided to the user as encapsulated new GIS operations.

3D Visualisation and Interaction

Visualisation is a general term to denote the process of extracting data from the model and representing them on the screen. This thesis concentrates on the representation of spatial data and, therefore, visualisation will be mostly used in the context of display of 3D graphics. Interaction with the 3D graphics on the screen is, in the broad sense (details are given below), detecting user actions and reaching them by re-computing the parameters of the models and producing of a new display. Since the 3D visualisation and interaction are critical elements the user interface in a 3D GIS, some fundamental principles (relevant for the thesis) are presented here.

The process of visualisation and interaction is completed in two steps, which are known as *pipelines,* i.e. input and output. The output pipeline comprises the process of sending information (primitives and attributes) to the screen; the input pipeline comprises the detection of user actions and the corresponding post-processing of the model. The visualisation schema is common for any program, with a graphics output, i.e. CAD systems, e.g. AutoCAD, Intergraph, Microstation, GIS programs, e.g. ArcVIEW, MapInfo, Web browsers, e.g. Netscape, Mosaic. The *application program* is the software controlling and ensuring the correct flow of data in both directions, i.e. 1) it accesses the data organised in a model, extracts a subset of data, which has to be on the screen at one particular moment and applying an appropriate graphics package (rendering engine), performs the necessary algorithms to transform 3D graphics to 2D image on the screen, and 2) it controls the input devices. The process of creating images from the model is often referred to as *rendering*.

The *graphics package* is a software interface to the graphics hardware, which provides means for the rendering, maintaining windows and detecting user actions. The tendency over the last few years to develop hardware-independent software has resulted in a *rendering package*, i.e. OpenGL. From a programming point of view, OpenGL is a set of standard procedures (written in C++) that facilitates the management of graphics hardware. To be able to run on different platforms, the OpenGL syntax does not support procedures dealing with windows and input devices. An application program using OpenGL has to utilise the windowing mechanism and the device tracking (e.g. detecting mouse click) of the current operation system (e.g. Windows, Unix), i.e. make use of corresponding libraries (e.g. Motif X for Unix). In return, OpenGL provides the means: to describe objects, arrange them in three-dimensional space, calculate the colour and illumination parameters and convert the mathematical description of the model to the pixel parameters on the screen. Similar rendering packages, but being less hardware independent, are Direct 3D (Microsoft) and Java 3D (Sun Microsystems), which is the newest 3D interface currently available for two operation systems, i.e. Windows and Solaris.

The visualisation of 3D graphics differs significantly from 2D graphics on several points: 1) scene creation, 2) tools for interaction and manipulation and 3) amount of data. The following text is meant to provide basic knowledge about the mechanisms to obtain readable scenes and ensure sufficient interaction tools.

Database Model

A database model is a theory or specification describing how a database is structured and used. Several such models have been suggested. Common models include:

- Flat model: This may not strictly qualify as a data model. The flat (or table) model consists of a single, two-dimensional array of data elements, where all members of a given column are assumed to be similar values, and all members of a row are assumed to be related to one another.
- Hierarchical model: In this model data is organized into a tree-like structure, implying a single upward link in each record to describe the nesting, and a sort field to keep the records in a particular order in each same-level list.
- Network model: This model organizes data using two fundamental constructs, called records and sets. Records contain fields, and sets define one-to-many relationships between records: one owner, many members.
- Relational model: is a database model based on first-order predicate logic. Its core idea is to describe a database as a collection of predicates over a finite set of predicate variables, describing constraints on the possible values and combinations of values.
- Object-relational model: Similar to a relational database model, but objects, classes and inheritance are directly supported in database schemas and in the query language.
- Star schema is the simplest style of data warehouse schema. The star schema consists of a few "fact tables" (possibly only one, justifying the name) referencing any number of "dimension tables". The star schema is considered an important special case of the snowflake schema.

Data Structure Diagram

A data structure diagram (DSD) is a diagram and data model used to describe conceptual data models by providing graphical notations which document entities and their relationships, and the constraints that binds them. The basic graphic elements of DSDs are boxes, representing entities, and arrows, representing relationships. Data structure diagrams are most useful for documenting complex data entities.

Data structure diagrams are an extension of the entity-relationship model (ER model). In DSDs, attributes are specified inside the entity boxes rather than outside of them, while relationships are drawn as boxes composed of attributes which specify the constraints that bind entities together. The E-R model, while robust, doesn't provide a way to specify the constraints between relationships, and becomes visually cumbersome when representing entities with several attributes. DSDs differ from the ER model in that the ER model focuses on the relationships between different entities, whereas DSDs focus on the relationships of the elements within an entity and enable users to fully see the links and relationships between each entity. There are several styles for representing data structure diagrams, with the notable difference in the manner of defining cardinality. The choices are between arrow heads, inverted arrow heads (crow's feet), or numerical representation of the cardinality.

Entity-relationship Model

An entity-relationship model (ERM) is an abstract conceptual data model (or semantic data model) used in software engineering to represent structured data. There are several notations used for ERMs.

Geographic Data Model

A data model in Geographic information systems is a mathematical construct for representing geographic objects or surfaces as data. For example,

- the vector data model represents geography as collections of points, lines, and polygons;
- the raster data model represent geography as cell matrixes that store numeric values;

- and the Triangulated irregular network (TIN) data model represents geography as sets of contiguous, nonoverlapping triangles.

Generic Data Model

Generic data models are generalizations of conventional data models. They define standardised general relation types, together with the kinds of things that may be related by such a relation type. Generic data models are developed as an approach to solve some shortcomings of conventional data models.

For example, different modelers usually produce different conventional data models of the same domain. This can lead to difficulty in bringing the models of different people together and is an obstacle for data exchange and data integration. Invariably, however, this difference is attributable to different levels of abstraction in the models and differences in the kinds of facts that can be instantiated (the semantic expression capabilities of the models).

The modelers need to communicate and agree on certain elements which are to be rendered more concretely, in order to make the differences less significant.

Semantic Data Model

A semantic data model in software engineering is a technique to define the meaning of data within the context of its interrelationships with other data. A semantic data model is an abstraction which defines how the stored symbols relate to the real world. A semantic data model is sometimes called a conceptual data model.

The logical data structure of a database management system (DBMS), whether hierarchical, network, or relational, cannot totally satisfy the requirements for a conceptual definition of data because it is limited in scope and biased toward the implementation strategy employed by the DBMS. Therefore, the need to define data from a conceptual view has led to the development of semantic data modelling techniques. That is, techniques to define the meaning of data within the context of its interrelationships with other data.The real world, in terms of resources, ideas, events, etc., are symbolically defined within physical data stores. A semantic data model is an abstraction which defines how the stored symbols relate to the real world. Thus, the model must be a true representation of the real world.

Data Model Topics

Data Architecture

Data architecture is the design of data for use in defining the target state and the subsequent planning needed to hit the target state. It is usually one of several architecture domains that form the pillars of an enterprise architecture or solution architecture. A data architecture describes the data structures used by a business and/or its applications. There are descriptions of data in storage and data in motion; descriptions of data stores, data groups and data items; and mappings of those data artifacts to data qualities, applications, locations etc. Essential to realizing the target state, Data architecture describes how data is processed, stored, and utilized in a given system. It provides criteria for data processing operations that make it possible to design data flows and also control the flow of data in the system.

Data Modelling

Data modelling in software engineering is the process of creating a data model by applying formal data model descriptions using data modelling techniques. Data modelling is a technique for defining business requirements for a database. It is sometimes called *database modelling* because a data model is eventually implemented in a database. The way data models are developed and used today. A conceptual data model is developed based on the data requirements for the application that is being developed, perhaps in the context of an activity model. The data model will normally consist of entity types, attributes, relationships, integrity rules, and the definitions of those objects. This is then used as the start point for interface or database design.

Data Properties

Some important properties of data for which requirements need to be met are:

- definition-related properties
  - o *relevance*: the usefulness of the data in the context of your business.
  - o *clarity*: the availability of a clear and shared definition for the data.
  - o *consistency*: the compatibility of the same type of data from different sources.
- content-related properties
  - o *timeliness*: the availability of data at the time required and how up to date that data is.
  - o *accuracy*: how close to the truth the data is.
- properties related to both definition and content
  - o *completeness*: how much of the required data is available.
  - o *accessibility*: where, how, and to whom the data is available or not available (e.g. security).
  - o *cost*: the cost incurred in obtaining the data, and making it available for use.

Data Organization

Another kind of data model describes how to organize data using a database management system or other data management technology. It describes, for example, relational tables and columns or object-oriented classes and attributes. Such a data model is sometimes referred to as the *physical data model*, but in the original ANSI three schema architecture, it is called "logical". In that architecture, the physical model describes the storage media (cylinders, tracks, and tablespaces). Ideally, this model is derived from the more conceptual data model described. It may differ, however, to account for constraints like processing capacity and usage patterns.

While *data analysis* is a common term for data modelling, the activity actually has more in common with the ideas and methods of *synthesis* (inferring general concepts from particular instances) than it does with *analysis* (identifying component concepts from more general ones). {*Presumably we call ourselves systems analysts because no one can say systems synthesists.*} Data modelling strives to bring the data structures of interest together into a cohesive, inseparable, whole by eliminating unnecessary data redundancies and by relating data structures with relationships.

A different approach is through the use of adaptive systems such as artificial neural networks that can autonomously create implicit models of data.

Data Structure

A data structure is a way of storing data in a computer so that it can be used efficiently. It is an organization of mathematical and logical concepts of data. Often a carefully chosen data structure will allow the most efficient algorithm to be used.

The choice of the data structure often begins from the choice of an abstract data type. A data model describes the structure of the data within a given domain and, by implication, the underlying structure of that domain itself.

This means that a data model in fact specifies a dedicated *grammar* for a dedicated artificial language for that domain. A data model represents classes of entities (kinds of things) about which a company wishes to hold information, the attributes of that information, and relationships among those entities and (often implicit) relationships among those attributes.

The model describes the organization of the data to some extent irrespective of how data might be represented in a computer system.

The entities represented by a data model can be the tangible entities, but models that include such concrete entity classes tend to change over time. Robust data models often identify abstractions of such entities. For example, a data model might include an entity class called "Person", representing all the people who interact with an organization. Such an abstract entity class is typically more appropriate than ones called "Vendor" or "Employee", which identify specific roles played by those people.

Data Model Theory

The term data model can have two meanings:

1. A data model *theory*, i.e. a formal description of how data may be structured and accessed.
2. A data model *instance*, i.e. applying a data model *theory* to create a practical data model *instance* for some particular application.

A data model theory has three main components:

- The structural part: a collection of data structures which are used to create databases representing the entities or objects modeled by the database.
- The integrity part: a collection of rules governing the constraints placed on these data structures to ensure structural integrity.
- The manipulation part: a collection of operators which can be applied to the data structures, to update and query the data contained in the database.

For example, in the relational model, the structural part is based on a modified concept of the mathematical relation; the integrity part is expressed in first-order logic and the manipulation part is expressed using the relational algebra, tuple calculus and domain calculus.

A data model instance is created by applying a data model theory. This is typically done to solve some business enterprise requirement. Business requirements are normally captured by a semantic logical data model. This is transformed into a physical data model instance from which is generated a physical database.

For example, a data modeler may use a data modelling tool to create an entity-relationship model of the corporate data repository of some business enterprise. This model is transformed into a relational model, which in turn generates a relational database.

Patterns

Patterns are common data modelling structures that occur in many data models.

SPATIAL AND NON-SPATIAL DATA MODELS

Spatial data sets are primarily defined as those which are directly or indirectly referenced to a location on the surface of the earth. When a dataset cannot be related to a location on the surface of the earth is referred as non spatial data. GIS technologies are unique in their capability to combine the two data sets thereby bringing a paradigm shift in thinking how the planning and monitoring system may work. The power and potential of such systems is unlimited thereby providing huge opportunity to process information which can be used effectively.

The non spatial data are numbers, characters or logical type. The spatial data sets, however has primary data type as point, line or polygon and may be referenced to some specific grid system. Traditionally the information systems in past have created the huge data repositories which appear to be non spatial in nature. However, these may indirectly be referring to specific locations.

Let us elaborate the discussion with example of a data table containing different types of soils and their characteristics. Here data are typically non spatial in nature as it directly or indirectly, does not refer to any location. In another example consider a table containing population information for specific locations say city

,districts or provinces. The population though is non spatial type, has relationship with locations. For large number of locations, the scope for use of such data could be to understand name of location having largest or least populations or statistical annotations eg, mean, average or other values for all the locations. If the location data is expressed specially incorporating latitude/longitude or linked to specific shape of locations eg districts or provinces on the map, the scope of analysis of data enhances many folds. Representation of population in defined slabs on map can be represented in different colours and would enable demarcation of areas of highest population or least population or identification of hot spots as per the requirements. The maps in such cases may be composed in many different ways. Such output are not only colourful and interesting to look at but also may accommodate display of large tabular data on one page. The Overlaying of different layers of spatial data such as transport, water or other natural resources may further provide an additional dimension for analysis of population variation. The following hence may be concluded;

Non-spatial Database Models

Motivation: Why database management systems

- Database management systems (DBMSs) are very good at organizing and managing large collections of persistent data.
  - o We use DBMSs to help cope with large amounts of data because, when problems get big, they get hard.
    - — Consider the task of finding a particular book in a typical university library.
    - — Now, reconsider that same task if the library doesnÂ't keep the books arranged in any particular order or if the library has no indexes.
  - o Using a big collection of unorganized things is practically impossible. Structure turns data into information.

  - o *Persistence* means that the data exist permanently; they do not disappear when the computer is shut off.
- DBMSs are like suitcases: they are somewhere to put stuff so that it's all in one place and easy to get to.
- DBMSs help protect data from unauthorized access.
- DBMSs help protect data from accidental corruption or loss due to:
  - o hardware failures such as power outages and computer crashes
  - o software failures such as operating system crashes
- DBMSs allow concurrent access, meaning that a single data set can be accessed by more than one user at a time
  - o virtually all commercial database applications require the data entry staff to have access to the database simultaneously.
    - — For example, an airline reservation system cannot restrict access to the database to a single travel agent.
  - o concurrent data access introduces unwanted problems caused by two users manipulating exactly the same data at exactly the same time.
    - — These problems can cause the database to be corrupted or for a user's interface program to never complete its query.
    - — These problems are analogous to road intersections: if there are no traffic lights or stop signs, havoc will ensue.
  - o DBMSs provide mechanisms to prevent concurrent access problems; these mechanisms are collectively called *concurrency control*.
- A *distributed* DBMS allows a single database to be split apart such that its pieces reside at geographically separated sites.
  - o this can provide performance improvements by eliminating transmitting the data across a relatively

slow long distance communication channel (it's a lot faster to have the database on your hard drive than to access it across an Ethernet or via a modem)

- o this can reduce concurrency control bottlenecks by giving each user that part of the database which they need rather than having all the users compete for access to the whole database

- DBMSs are not necessarily meant for data analysis; that is more the job of a spread sheet or some other special-purpose analysis tool.
  - o DBMSs are general-purpose tools. It is basically irrelevant to the DBMS what is stored within it. Software design principles suggest de-coupling domain specific analysis packages from the DBMS to keep the division of labor clear.
  - o DBMSs are very good at retrieving a relatively small portion of the database and passing it along for detailed analysis by a tool designed for that purpose.
  - o DBMSs often allow integrity constraints to be imposed on the data to insure validity and consistency. These rules can interfere with ad-hoc analysis in which the user manipulates the data without any preconceived ideas of how the data should relate to each other.
  - o DBMSs often do not have adequate facilities to perform complicated calculations; some have no such facilities whatsoever.

Data Flow Diagram

A data flow diagram (DFD) is a graphical representation of the "flow" of data through an information system. It differs from the flowchart as it shows the *data* flow instead of the *control* flow of the program. A data flow diagram can also be used for the visualization of data processing. Data flow diagrams were invented by Larry Constantine, the original developer of structured design, based on Martin and Estrin's "data flow graph" model of computation. It is common practice to draw a context-level Data flow diagram first which shows the interaction between the system and outside entities. The DFD is designed to show how a system is divided into smaller portions and to highlight the flow of data between those parts. This context-level Data flow diagram is then "exploded" to show more detail of the system being modeled.

Information Model

An Information model is not a type of data model, but more or less an alternative model. Within the field of software engineering both a data model and an information model can be abstract, formal representations of entity types that includes their properties, relationships and the operations that can be performed on them. The entity types in the model may be kinds of real-world objects, such as devices in a network, or they may themselves be abstract, such as for the entities used in a billing system. Typically, they are used to model a constrained domain that can be described by a closed set of entity types, properties, relationships and operations.

According to Lee (1999) an information model is a representation of concepts, relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse. It can provide sharable, stable, and organized structure of information requirements for the domain context. More in general the term *information model* is used for models of individual things, such as facilities, buildings, process plants, etc.

In those cases the concept is specialised to Facility Information Model, Building Information Model, Plant Information Model, etc. Such an information model is an integration of a model of the facility with the data and documents about the facility.

An information model provides formalism to the description of a problem domain without constraining how that description is mapped to an actual implementation in software.

There may be many mappings of the information model. Such mappings are called data models, irrespective of whether they are object models (e.g. using UML), entity relationship models or XML schemas.

Object Model

An object model in computer science is a collection of objects or classes through which a program can examine and manipulate some specific parts of its world. In other words, the object-oriented interface to some service or system. Such an interface is said to be the *object model of* the represented service or system.

For example, the Document Object Model (DOM) is a collection of objects that represent a page in a web browser, used by script programs to examine and dynamically change the page.

There is a Microsoft Excel object model for controlling Microsoft Excel from another program, and the ASCOM Telescope Driver is an object model for controlling an astronomical telescope.

In computing the term *object model* has a distinct second meaning of the general properties of objects in a specific computer programming language, technology, notation or methodology that uses them. For example, the *Java object model*, the *COM object model*, or *the object model of OMT*. Such object models are usually defined using concepts such as class, message, inheritance, polymorphism, and encapsulation. There is an extensive literature on formalized object models as a subset of the formal semantics of programming languages.

Object-Role Model

Object-Role Modelling (ORM) is a method for conceptual modelling, and can be used as a tool for information and rules analysis.

Object-Role Modelling is a fact-oriented method for performing systems analysis at the conceptual level. The quality of a database application depends critically on its design.

To help ensure correctness, clarity, adaptability and productivity, information systems are best specified first at the conceptual level, using concepts and language that people can readily understand. The conceptual design may include data, process and behavioral perspectives, and the actual DBMS used to implement the design might be based on one of many logical data models (relational, hierarchic, network, object-oriented etc.).

Unified Modelling Language Models

The Unified Modelling Language (UML) is a standardized general-purpose modelling language in the field of software engineering. It is a graphical language for visualizing, specifying, constructing, and documenting the artifacts of a software-intensive system. The Unified Modelling Language offers a standard way to write a system's blueprints, including:

- Conceptual things such as business processes and system functions
- Concrete things such as programming language statements, database schemas, and
- Reusable software components.

UML offers a mix of functional models, data models, and database models.

3

Data Visualization

INTRODUCTION

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions.

Data visualization is a general term that describes any effort to help people understand the significance of data by placing it in a visual context. Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognized easier with data visualization software.

Today's data visualization tools go beyond the standard charts and graphs used in Microsoft Excel spreadsheets, displaying data in more sophisticated ways such as infographics, dials and gauges, geographic maps, sparklines, heat maps, and detailed bar, pie and fever charts.

The images may include interactive capabilities, enabling users to manipulate them or drill into the data for querying and analysis. Indicators designed to alert users when data has been updated or predefined conditions occur can also be included.

Importance of data visualization

Data visualization has become the de facto standard for modern business intelligence (BI). The success of the two leading vendors in the BI space, Tableau and Qlik -- both of which heavily emphasize visualization -- has moved other vendors toward a more visual approach in their software. Virtually all BI software has strong data visualization functionality.

Data visualization tools have been important in democratizing data and analytics and making data-driven insights available to workers throughout an organization. They are typically easier to operate than traditional statistical analysis software or earlier versions of BI software. This has led to a rise in lines of business implementing data visualization tools on their own, without support from IT.

Data visualization software also plays an important role in big data and advanced analytics projects. As businesses accumulated massive troves of data during the early years of the big data trend, they needed a way to quickly and easily get an overview of their data. Visualization tools were a natural fit.

Visualization is central to advanced analytics for similar reasons. When a data scientist is writing advanced predictive analytics or machine learning algorithms, it becomes important to visualize the outputs to monitor results and ensure that models are performing as intended. This is because visualizations of complex algorithms are generally easier to interpret than numerical outputs.

Examples of data visualization

Data visualization tools can be used in a variety of ways. The most common use today is as a BI reporting tool. Users can set up visualization tools to generate automatic dashboards that track company performance across key performance indicators and visually interpret the results.

Many business departments implement data visualization software to track their own initiatives. For example, a marketing team might implement the software to monitor the performance of an email campaign, tracking metrics like open rate, click-through rate and conversion rate.

As data visualization vendors extend the functionality of these tools, they are increasingly being used as front ends for more sophisticated big data environments. In this setting, data visualization software helps data engineers and scientists keep track of data sources and do basic exploratory analysis of data sets prior to or after more detailed advanced analyses.

How data visualization works

Most of today's data visualization tools come with connectors to popular data sources, including the most common relational databases, Hadoop and a variety of cloud storage platforms. The visualization software pulls in data from these sources and applies a graphic type to the data.

Data visualization software allows the user to select the best way of presenting the data, but, increasingly, software automates this step. Some tools automatically interpret the shape of the data and detect correlations between certain variables and then place these discoveries into the chart type that the software determines is optimal.

Typically, data visualization software has a dashboard component that allows users to pull multiple visualizations of analyses into a single interface, generally a web portal.

Graphic Display Techniques

Traditional maps are abstractions of the real world, a sampling of important elements portrayed on a sheet of paper with symbols to represent physical objects. People who use maps must interpret these symbols. Topographic maps show the shape of land surface with contour lines or with shaded relief. Today, graphic display techniques such as shading based on altitude in a GIS can make relationships among map elements visible, heightening one's ability to extract and analyse information. For example, two types of data were combined in a GIS to produce a perspective view of a portion of San Mateo County, California.

- The digital elevation model, consisting of surface elevations recorded on a 30-meter horizontal grid, shows high elevations as white and low elevation as black.
- The accompanying Landsat Thematic Mapper image shows a false-colour infrared image looking down at the same area in 30-meter pixels, or picture elements, for the same coordinate points, pixel by pixel, as the elevation information.

A GIS was used to register and combine the two images to render the three-dimensional perspective view looking down the San Andreas Fault, using the Thematic Mapper image pixels, but shaded using the elevation of the landforms. The GIS display depends on the viewing point of the observer and time of day of the display, to properly render the shadows created by the sun's rays at that latitude, longitude, and time of day. An archeochrome is a new way of displaying spatial data. It is a thematic on a 3D map that is applied to a specific building or a part of a building. It is suited to the visual display of heat loss data.

Spatial ETL

Spatial ETL tools provide the data processing functionality of traditional Extract, Transform, Load (ETL) software, but with a primary focus on the ability to manage spatial data. They provide GIS users with the ability to translate data between different standards and proprietary formats, whilst geometrically transforming the data en-route.

GIS Developments

Many disciplines can benefit from GIS technology. An active GIS market has resulted in lower costs and continual improvements in the hardware and software components of GIS. These developments will, in turn, result in a much wider use of the technology throughout science, government, business, and industry, with applications including real estate, public health, crime mapping, national defence, sustainable development, natural resources, landscape architecture, archaeology, regional and community planning, transportation and logistics. GIS is also diverging into location-based services (LBS). LBS allows GPS enabled mobile devices to display their location in relation to fixed assets (nearest restaurant, gas station, fire hydrant), mobile assets (friends, children, police car) or to relay their position back to a central server for display or other processing. These services continue to develop with the increased integration of GPS functionality with increasingly powerful mobile electronics (cell phones, PDAs, laptops).

OGC Standards

The Open Geospatial Consortium (OGC) is an international industry consortium of 384 companies, government agencies, universities and individuals participating in a consensus process to develop publicly available geoprocessing specifications. Open interfaces and protocols defined by OpenGIS Specifications support interoperable solutions that "geo-enable" the Web, wireless and location-based services, and mainstream IT, and empower technology developers to make complex spatial information and services accessible and useful with all kinds of applications. Open Geospatial Consortium (OGC) protocols include Web Map Service (WMS) and Web Feature Service (WFS).

GIS products are broken down by the OGC into two categories, based on how completely and accurately the software follows the OGC specifications. *Compliant Products* are software products that comply to OGC's OpenGIS Specifications. When a product has been tested and certified as compliant through the OGC Testing Program, the product is automatically registered as "compliant" on this site. *Implementing Products* are software products that implement OpenGIS Specifications but have not yet passed a compliance test. Compliance tests are not available for all specifications. Developers can register their products as implementing draft or approved specifications, though OGC reserves the right to review and verify each entry.

The advantages and benefits of good data visualization

Data visualization is another form of visual art that grabs our interest and keeps our eyes on the message. When we see a chart, we quickly see trends and outliers. If we can see something, we internalize it quickly. It's storytelling with a purpose. If you've ever stared at a massive spreadsheet of data and couldn't see a trend, you know how much more effective a visualization can be.

TYPES OF VISUALIZATION

The different types of visualizations

When you think of data visualization, your first thought probably immediately goes to simple bar graphs or pie charts. While these may be an integral part of visualizing data and a common baseline for many data graphics, the right visualization must be paired with the right set of information. Simple graphs are only the tip of the iceberg. There's a whole selection of visualization methods to present data in effective and interesting ways.

Common general types of data visualization:

- Charts
- Tables
- Graphs
- Maps
- Infographics
- Dashboards

More specific examples of methods to visualize data:

- Area Chart
- Bar Chart
- Box-and-whisker Plots
- Bubble Cloud
- Bullet Graph

- Cartogram
- Circle View
- Dot Distribution Map
- Gantt Chart
- Heat Map
- Highlight Table
- Histogram
- Matrix
- Network
- Polar Area
- Radial Tree
- Scatter Plot (2D or 3D)
- Streamgraph
- Text Tables
- Timeline
- Treemap
- Wedge Stack Graph
- Word Cloud
- And any mix-and-match combination in a dashboard!

GIS and Visualization

Visualizing large amounts of information interactively is one of the most attractive and useful capabilities of GIS. High-powered computers can alter any element of the display "on the fly," changing not only the look of the graphic image but also its interpretation. This ability to create multiple perspectives -- both literally and figuratively -- enhances a viewer's perceptive abilities to understand the phenomenon being studied like never before.

Visualizing data using current computing technology and interactive GIS has many advantages over doing so using traditional paper maps. The following is a partial list, one that grows with each new version of software and each new advance in hardware.

- GIS is fully interactive. Adding new fields of data, taking them off, changing the color scheme or form of the map, adding text, moving symbols, and a host of other capabilities give a user tremendous flexibility and power.
- GIS displays are zoomable and pannable. Moving around in the display offers a user new perspectives, greater (or less) detail, and new insight.
- Users can take advantage of computationally intensive functions such as "draping" a perspective view over a surface (like a digital elevation model) or creating the impression of three dimensions on a 2D display (the computer screen) using complex rendering and shading algorithms.

Many of the tips and guidelines outlined in this unit were developed by cartographers over many decades for the design of paper maps. They are all relevant to the design of interactive displays within GIS. In fact, with the vast array of choices facing a GIS display designer and with the luxury of complete interactivity, awareness of the power of the visual image to persuade a viewer has never been more important.

Visualizing Temporal Data with GIS

The visualization of temporal data is important for a variety of sciences and disciplines in understanding space-time factors and change. Analytically, showing change in a setting is important for decision-making and forecasting future events. Different sets of software and techniques have been developed recently to apply spatiotemporal visualization and analysis.

Open source tools have recently developed methods to process open formated files such as Keyhole Markup Language (KML) files. For instance, the R statistical package can use the plotKML() package to create KML files that can then be used to visualize spatiotemporal change within the XML-like data in a given area of analysis. The parsing order and user can view templates that can even be specified to display the visual output.

In the social sciences, using social media has become popular in research. Data can be geolocated and be represented in spatial cubes that encapsulate location data and unstructured information such as texts and photographs. These data can be assessed across different time intervals to demonstrate given spatial patterns of movement or interactions using social media streams. This development has been dubbed as part of CyberGIS, where it can be used, among other areas, to allow an understanding and study of social patterns as they are affected by events, such as sickness (e.g., such as a flu pandemic). In effect, these spatial cubes that are created can be useful for emergency responses to monitor how mass events could affect regions so that officials can best respond to real-time events.

A flow map of number of travels made by potential flu-affected Twitter users during seven days (January 29th to February 5th, 2013) from the Los Angeles to the other areas of North America. Source: Cao et al., 2015.

One aspect of research that has seen a lot of interest is network analysis. In order to understand interaction across space and time, using networks of entities (e.g., people, fauna, etc.) interacting and determining the importance of interactions has become one focus area. Geotagged data can be imported using ArcScene, as one example, where network change and interactions are determined by looking at how information diffuses from a source onto other actors or individuals in a given networked space. In effect, this allows the measurement of the perceived importance of events on the actors themselves and how those events transpire across a network of actors.

Issues of access of spatiotemporal data have been indicated as a drawback and limitation in countries where data are needed for analysis but more difficult to obtain. This can have planning and management repercussion, for instance, in urban development. Countries that are rapidly growing in population and urbanism are facing some of the greatest change to their societies, but these countries often do not have resources to make informed decisions.

One project, MEGA-WEB Geo Webservices, has developed online spatiotemporal visualization and analysis to show remote sensing and other collected data. Having a platform that can visualize multiple datasets, including remote sensing data of urban growth, facilitates spatiotemporal assessment without having to access tools that often have expensive licensing fees.

The complexity of spatiotemporal data is often evident in querying multidimensional data across space and time. Tools such as the SubVizCon framework have been developed to allow users easier access to specify subset multidimensional queries. This tool can be integrate within ArcGIS to visualize the data and is useful for investigating patterns within such data.

While spatiotemporal data and visualization have remained a challenge, sometimes due to the complexity of the data, increasingly we are seeing tools that are available on common open source platforms (e.g., R statistical package) or commercial platforms such as ArcGIS. Both the social and natural sciences, in particular, have benefited and have created tools that are potentially informative and useful for decision-making. The use of spatiotemporal visualization is likely to be more significant as major global change will require real-time data integration and visual analysis.

MAP LAYOUT DESIGN AND SYMBOLOGY

Map layout

Map Layout is the assembling of the various elements of a map into a single whole, including the map itself, its legend, title, scale bars, and other elements. Map symbolism can rarely stand alone to sufficiently depict all the necessary information that a map is trying to tell; additional explanation and context is usually needed. Their primary purpose is to give a place identity, orientation, subject matter, symbolization, etc. This term usually refers to the combination of the map image with auxiliary elements; the assembling of the geographic symbols within the map image is called Map composition.

Legends

A map legend, or map key, is an explanation of the symbols or pictorial language and convention of the map. Legend content, design and placement are important because a map can fail to communicate its message if readers cannot determine what the symbols in the map are intended to represent. [1]" Not all symbols used on the map necessarily need to be included in the legend; Symbols that are intuitive, conventional, well-labeled, or unimportant may be left out with careful consideration. Maps with highly standardized symbology, such as street maps or topographic maps, may remove the legend altogether.

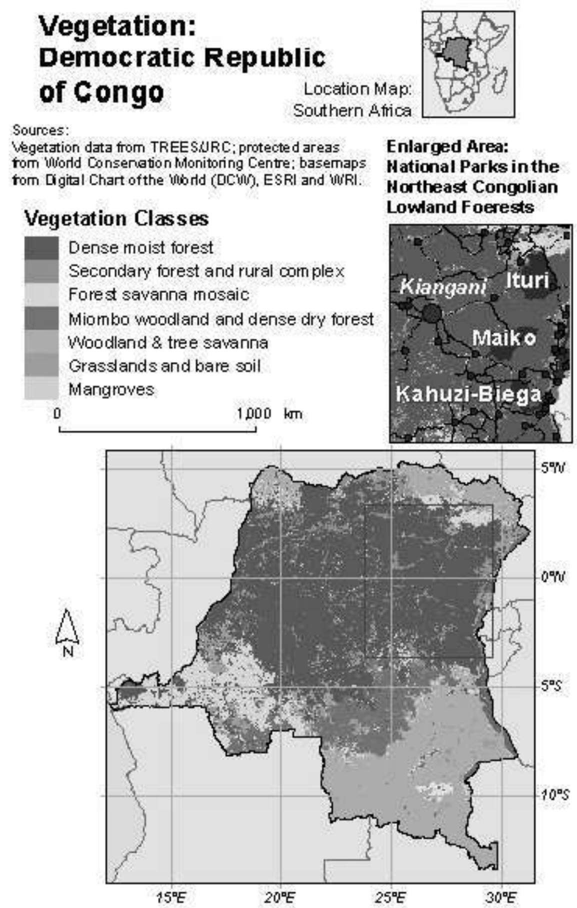
Elements of legend design

- A legend includes four parts: content, wording, placement, and style.
- Symbols must look exactly like the symbol used on the map in order to ensure viewer comprehension. This is particularly important in regards to size.
- The legend title should be explanatory and make a clear connection to the map's legend.
- Type used in the legend does not have to match type used on the map, although the typefaces should work well together and maintain clarity.
- Placement of the legend should be based on balance and white space.

Layout and Symbolization

Designing a Layout

When you design a map, you need to carefully consider the relative balance of all of its visual elements.



You have the map itself, and perhaps you have multiple maps (an inset showing where the location of interest is in relation to the rest of the world, for example). You'll also probably have a title, a legend, scale bar, source information, and other little bits and pieces. Map nerds (cartographers) call those bits and pieces (aside from the main map) *marginalia*.

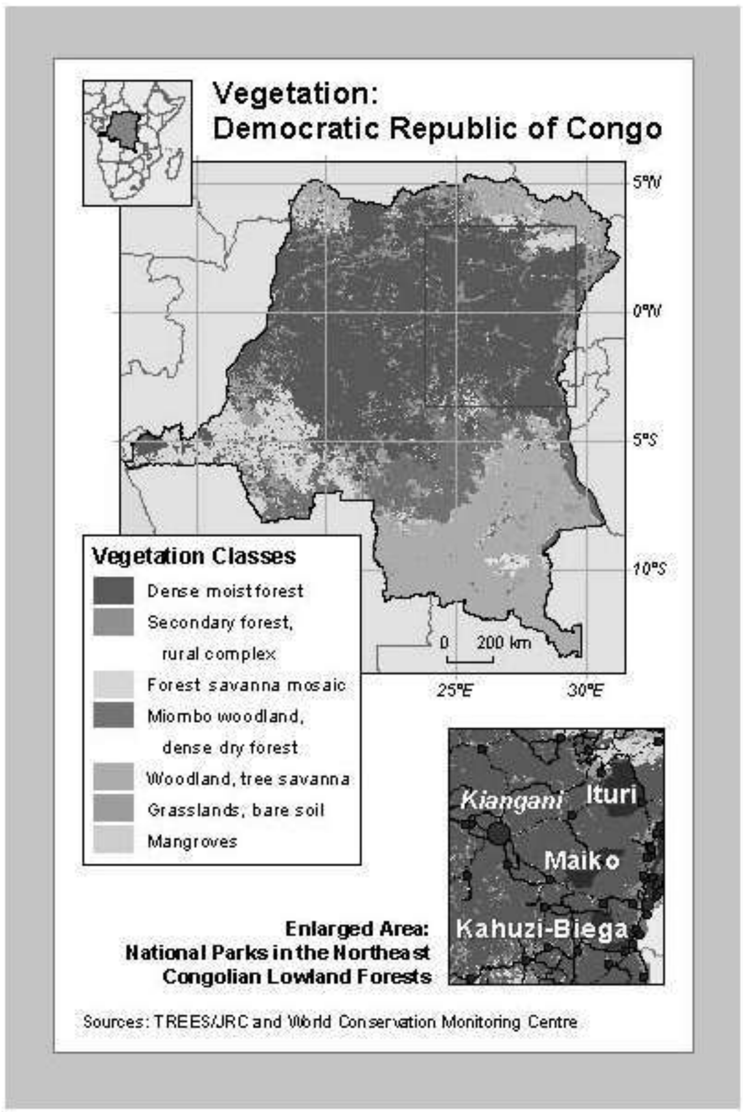


Figure: Good Map Styling Vs Poor Map Styling

The act of sorting things out so that clarity is maximized is called establishing a *visual hierarchy*. Consider the two examples below from Designing Better Maps: A Guide for GIS Users by Cindy Brewer to see what this means in practice. It's all about shaping the layout items around your message. The map on the top has all of the necessary elements, but they're not balanced on the page. The example on the bottom shows the same elements arranged in a way that promotes clarity and harmony in the map design. The first example is a *bad* layout. The second one is a *good* layout.

Symbolization

By this point you've encountered a wide range of ways in which things can be represented on maps. You've worked a lot with choropleth (colored-area) maps in labs and you've seen a lot of those as examples in the course content. There are many other ways to show stuff though, and I want you to be aware of some of the most common methods.

Point Symbols

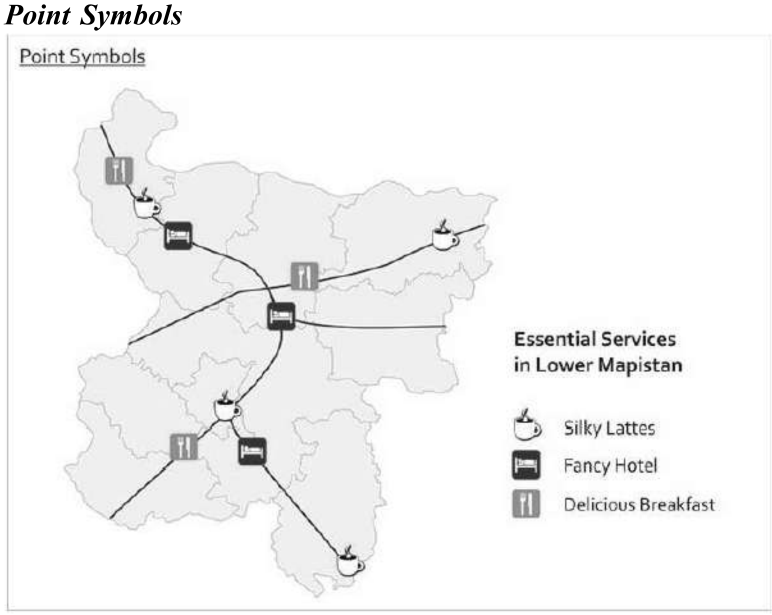


Figure: Map of "Lower Mapistan" with Point Symbols

One of the simplest things you can do is to use icons to represent point features. We call these point symbols, and they can range from very detailed in their design (you could even use small photographs if you really wanted to), to very abstract and iconic. Here's an example of point symbols used to show the locations of important services in Lower Mapistan. Good resources for point symbols include MapBox's Maki set and our own Penn State Symbol Store.

Proportional & Graduated Symbols

Another form of point symbol is one that changes its size based on underlying data values. These are called proportional symbols when you size each symbol *in relation to its attribute value*. In a proportional symbol map it's possible for every symbol to be a slightly different size, since size relates to the data value itself. In contrast, graduated symbols use *preset symbol sizes to represent a category of values*. It's usually easier for people to make quick comparisons with graduated symbols, while proportional symbols do a nicer job of revealing the underlying diversity in your dataset. You can see the difference here with this map of cable bill delinquency. Both methods are good to use if you want to map raw values rather than rates (remember our previous work on normalization?).

MAP ELEMENTS: SCALE, PROJECTION, COORDINATE SYSTEMS

Mapping Object Model to Data Model

Visual Paradigm supports Object Relational Mapping (ORM) which maps object models to entity relational models and vice versa. Visual Paradigm helps mapping between Java objects to relational database. It not only preserves the data, but also the state, foreign/primary key mapping, difference in data type and business logic. Thus, you are not required to handle those tedious tasks during software development.

Mapping classes to entities

Generally speaking, the mapping between class and entity is a one-to-one mapping, meaning that one class in object model maps with one entity in data model. Classes that map with entities are represented by the stereotype <<ORM Persistable>>.

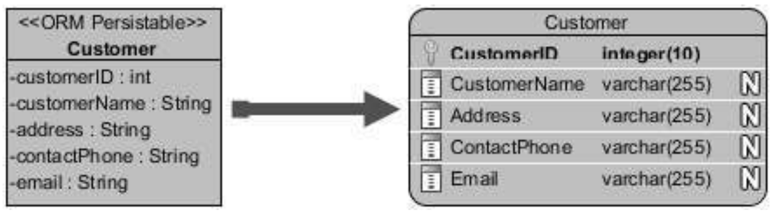


Fig. Mapping class

In the above example, the *Customer* class is the object equivalent of the *Customer* entity. This means that in application development or in runtime, an instance of *Customer* (class) stores the information of a customer retrieved from the *Customer* table of database.

Mapping attributes to columns

Since the persistent classes map to the entities, persistent attributes map to columns accordingly. Visual Paradigm ignores all non persistent attributes such as derived values.

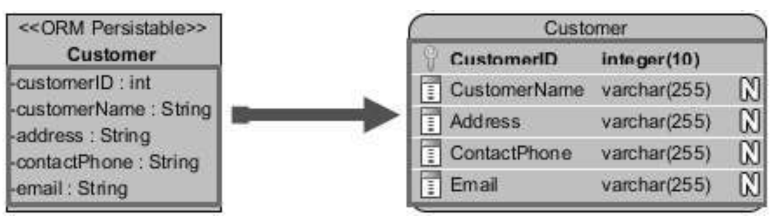


Fig. Mapping attributes

Mapping data type

Persistent attribute types are automatically mapped to appropriate column data types of the database you desired. The following table lists out the typical mapping between object model and data model. Note that the actual data type to map to depends on the default database you selected in database configuration.

Mapping primary key

You can map an attribute to a primary key column. When you synchronize the ORM-Persistable Class to the ERD, you will be prompted by a window to select primary key.

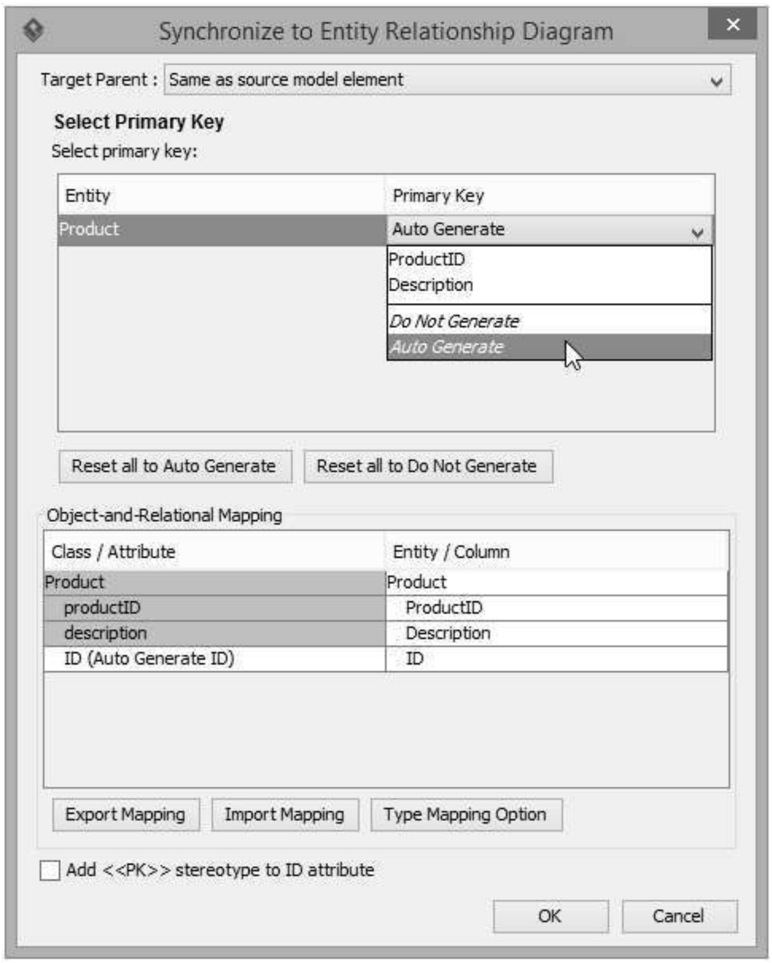


Fig. Selecting the way to map primary key

You can select an existing attribute as primary key, let us generate one for you. or select Do Not Generate to leave the generated entity without primary key.

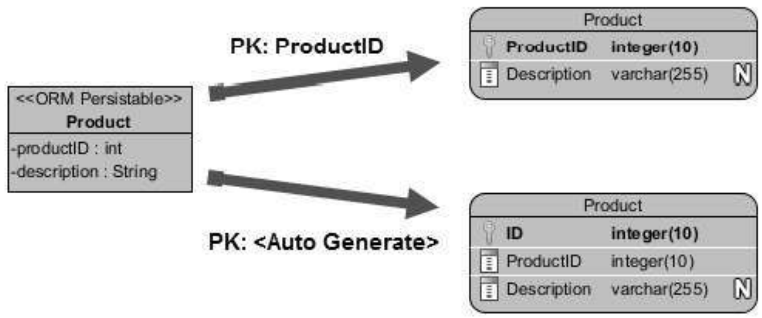


Fig. Mapping primary key

The above diagram shows if you assign *ProductID* as primary key, the *ProductID* of the generated entity, *Product* will become bold; whereas if you select Auto Generate for the primary key, Visual Paradigm generates an additional attribute*ID* as the primary key of the *Product* entity.

Mapping association

Association represents a binary relationship among classes. Each class of an association has a role. A role name is attached at the end of an association line. Visual Paradigm maps the role name to a phrase of relationship in the data model.

Mapping aggregation

Aggregation is a stronger form of association. It represents the "has-a" or "part-of" relationship.

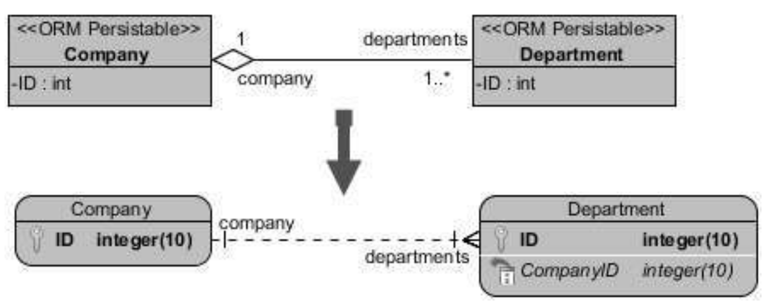


Fig. Mapping aggregation

In the above example, it shows that a company consists of one or more department while a department is a part of the company.

Coordinate Systems

Implicit with any GIS data is a spatial reference system. It can consist of a simple arbitrary reference system such as a 10 m x 10 m sampling grid in a wood lot or, the boundaries of a soccer field or, it can consist of a geographic reference system, i.e. one where the spatial features are mapped to an earth based reference system. The focus of this topic is on earth reference systems which can be based on a Geographic Coordinate System (GCS) or a Project Coordinate System (PCS).

Geographic Coordinate Systems

A geographic coordinate system is a reference system for identifying locations on the curved surface of the earth. Locations on the earth's surface are measured in angular units from the center of the earth relative to two planes: the plane defined by the equator and the plane defined by the prime meridian (which crosses Greenwich England). A location is therefore defined by two values: a latitudinal value and a longitudinal value.

A latitude measures the angle from the equatorial plane to the location on the earth's surface. A longitude measures the angle between the prime meridian plane and the north-south plane that intersects the location of interest. For example Colby College is located at around 45.56° North and 69.66° West. In a GIS system, the North-South and East-West directions are encoded as signs. North and East are assigned a positive (+) sign and South and West are assigned a negative (-) sign. Colby College's location is therefore encoded as +45.56° and -69.66°.

Projection

Maps are representations of the earth's surface. This representation of space requires a spatial reference system based upon a set of geometric assumptions. A spatial reference system establishes a point of origin, orientation of reference axes, and geometric meaning of measurements, as well as units of measure. This locational information describes the position of particular geographic features on the Earth's surface, as well as spatial relationships between features, or georeferencing.

Georeferencing is the process of establishing a relationship between the data displayed in your GIS software and its real-world location using a coordinate system. While you can locate a point on the Earth with great accuracy, representing the same point on a map is still an approximation.

Overview of Coordinate Systems, Datums, and Projections

Coordinate systems

Coordinates are sets of measurements related to a specific spatial reference system, e.g. pairs of distance measurements (X,Y) on independent planar reference systems, or to angular measurements from the plane of the equator and the Prime Meridian called latitude-longitude pairs.

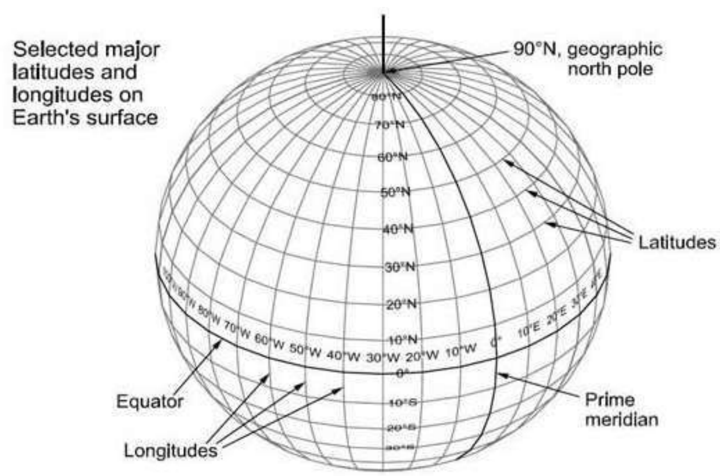


Fig. :† Global Coordinate System with main components labeled.†

Coordinate systems allow for georeferencing. Every location in our thematic data is referenced to its corresponding location on the Earth's surface. Coordinate systems are made up of an ellipsoid, datum, projection and units.

A common system of spatial reference is a critical element of a GIS project or map set, since it brings the different map layers into correspondence. There are two types of horizontal reference systems: geographic coordinate system and plane coordinate systems. The geographic coordinate system specifies locations on a spherical Earth; and a plane coordinate system specifies locations on a flattened Earth.

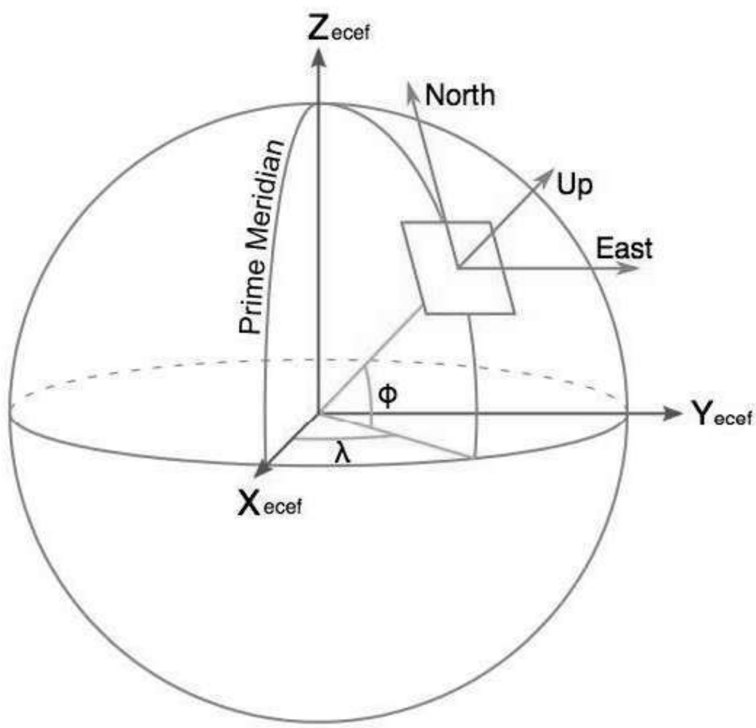


Fig. :† (Above) X, Y, and Z-axes (blue) from which latitude and longitude angles (orange) are measured. Green box shows how these angles interest and locate an area on the Earth's surface.

Latitude (parallels) and longitude (meridians) form an imaginary network over the earth's surface (graticule). Degrees of latitude and longitude (Global Reference System) are used to locate exact positions on the surface of the globe; they are not uniform measures on the Earth's surface. This rrence system measures angles from the center of the Earth, rather than distances on the Earth's surface. Additionally, the global coordinate system is used for the curved surface of the Earth - it is not a map projection. Only at the equator does the distance of one degree of longitude approximately equal the distance of one degree of latitude.

Degrees of latitude measure north and south of the equator (0 - 90 degrees). Latitude also represents an important measure of seasonality.

Degrees of longitude measure east and west of the Prime Meridian (0 - 180 degrees). The Prime Meridian has changed over time due to changing political fortunes and opportunities, but is currently based on Greenwich, England. Degrees of longitude also measure time, where 15 degrees of longitude is equal to one hour.

4

Introduction to GIS Software (Open Source)

Introduction to GIS

A Geographic Information System combines computer cartography with a database management system. The major components common to a GIS. This diagram suggests that a GIS consists of three subsystems:

(1) an input system that allows for the collection of data to be used and analysed for some purpose;

(2) computer hardware and software systems that store the data, allow for data management and analysis, and can be used to display data manipulations on a computer monitor;

(3) an output system that generates hard copy maps, images, and other types of output.

Two basic types of data are normally entered into a GIS. The first type of data consists of real world phenomena and features that have some kind of spatial dimension. Usually, these data elements are depicted mathematically in the GIS as either points, lines, or polygons that are referenced geographically (or geocoded) to some type of coordinate system. This type data is entered into the GIS by devices like scanners, digitizers, GPS, air photos, and satellite imagery. The other type of data is sometimes referred to as an attribute. Attributes are pieces of data that are connected or related to the points, lines, or polygons. This attribute data can be analysed to determine patterns of importance. Attribute data is entered directly into a database where it is associated with element data.

Plotted data can be defined as elements because their main purpose is to describe the location of the earthquakes.

For each of the earthquakes, the GIS also has data on their depth. These measurements can be defined as attribute data because they are connected to the plotted earthquake locations shows the attribute earthquake depth organized into three categories: shallow; intermediate; and deep. This analysis indicates a possible relationship between earthquake depth and spatial location-deep earthquakes do not occur at the mid-oceanic ridges.

Within the GIS database a user can enter, analyse, and manipulate data that is associated with some spatial element in the real world. The cartographic software of the GIS enables one to display the geographic information at any scale or projection and as a variety of layers which can be turned on or off. Each layer would show some different aspect of a place on the Earth. These layers could show things like a road network, topography, vegetation cover, streams and water bodies, or the distribution of annual precipitation received.

The difference between element and attribute data can be illustrated. The location of some of the earthquakes that have occurred in the last century. They plotted data points can be defined as elements because their main purpose is to describe the location of the earthquakes.

For each of the earthquakes, the GIS also has data on their depth. These measurements can be defined as attribute data because they are connected to the plotted earthquake locations. The attribute earthquake depth organized into three categories: shallow; intermediate; and deep. This analysis indicates a possible relationship between earthquake depth and spatial location-deep earthquakes do not occur at the mid-oceanic ridges. Within the

GIS database a user can enter, analyse, and manipulate data that is associated with some spatial element in the real world. The cartographic software of the GIS enables one to display the geographic information at any scale or projection and as a variety of layers which can be turned on or off.

Each layer would show some different aspect of a place on the Earth. These layers could show things like a road network, topography, vegetation cover, streams and water bodies, or the distribution of annual precipitation received. The merges data layers for vegetation community type, glaciers and ice fields, and water bodies (streams, lakes, and ocean).

What is a GIS?

Geographic Information Systems provide a method for integrating and analysing spatial (digital map based) information such as "where is the nearest movie theatre?" alongside related non-spatial information (what movies are playing there?). GIS have three major capabilities (computer mapping, spatial analysis and spatial database) and can operate on a range of platforms (desktop/laptop computer, Internet, PDA, etc). Many people are becoming far more familiar with seeing the results both textually-for example when their phone shows them the nearest pub-and on open map systems such as Google Maps. Where in the past people had to literally use pencils and string on a paper map to find their nearest school, a computer can do this now extremely quickly an accurately, as long as all the information has been entered correctly in the first place.

In a broader context, GIS involves people and often brings a philosophy of change. For example, in 1994, the New York Police Department introduced GIS to locate crime 'hot-spots', analyse underlying problems and devise strategies and solutions to deal with the problems. Since 1993, violent crime has dropped by two-thirds in New York City. This strategy, known as COMPSTAT, has expanded to cities and jurisdictions across the United States and around the world.

GIS Software

One leading GIS software vendor is ESRI, based in Redlands, California, which offers ArcGIS for the desktop, ArcGIS Server for Internet mapping, ArcPad for PDAs and a range of other products and services for developers. Other popular GIS software packages are available from Cadcorp, Intergraph, MapInfo, Manifold and Autodesk. ERDAS Imagine, ENVI, Idrisi, and PCI Geomatica are geared towards remote sensing i.e analysis of satellite/aircraft images.

There are many third-party extensions and utilities for ArcGIS and other GIS and raster software platforms. Currently, open source GIS software options can be chosen from the first OS GIS package GRASS, recent open source options are DIVA GIS, Quantum GIS, and uDig. There are efforts underway, through the Open GIS Consortium to provide interoperability among spatial data formats and software. The leading contender for spatial data storage is another open source package called PostGIS, which is a spatial extension to the open source database PostgresSQL.

Why is Geospatial Topology Important In A GIS?

Geospatial analysis provides a unique perspective on the world, a tool through which to examine events, patterns, and processes that operate on or near the surface of our planet. To have this tool interact also with the relevant topological information is a must, that is needed to have the complete understanding of the way rivers flow, the weather is affected by the terrain or even how humans and other animals flow or chose to a specific habitat, this kind of information can only be gathered and useful if topological information is present in the GIS.

DATA INPUTS SCANNING/ACQUIRING DATA

Data Input 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.

The choice of data input method is governed largely by the application, the available budget, and the type and the complexity of data being input.

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.

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.

Spatial Data

Spatial data comes in two major formats, as vector or raster data. The main difference is that a raster is usually a static background picture used to illustrate, whereas a vector is an intelligent ladder of information that can be selected and searched.

Vector (points, lines, polygons)

Vector data often represents anthropogenic (human) features such as roads, buildings, political boundaries (counties, congressional districts, etc), and other features such as lakes and rivers. Vector data is scalable without loss of resolution and is generally represented by XYZ points in a Cartesian frame reference.

Raster (grids/images)

Raster data is pixellated data, and the more pixels that map the data, the better the resolution. However, if raster data is enlarged, it simply enlarges the pixels, which then leads to a loss of resolution. There are many moves being made to make raster data more useable/searchable as it is much faster to collect, unlike vectors where each piece of data usually has a manual input origin.

Raster data is usually derived from satellite imagery or aerial photography (known as remote sensing). Ordinary cameras are only sensitive to visible light. Satellite sensors can capture not only visible light, but also the thermal, microwave, infrared or other types of energy emanating from the Earth's surface. This extra data provides information about sea surface temperatures, vegetation, ozone, etc. Remote sensing is also used to study other planets and extraterrestrial bodies, such as Mars.

GEOREFERENCING OF MAPS

A common forestry task would be the update of the information for a forestry area. It is possible that the previous information for that area dates several years back and was collected analogically (that is, in paper) or perhaps it was digitized but all you have left is the paper version of that inventory data.

Most likely you would like to use that information in your GIS to, for example, compare later with later inventories. This means that you will need to digitize the information at hand using your GIS software. But before you can start the digitizing, there is an important first step to be done, scanning and georeferencing your paper map.

Scan the map

The first task you will have to do is to scan your map. If your map is too big, then you can scan it in different parts but keep in mind that you will have to repeat preprocessing and georeferencing tasks for each part. So if possible, scan the map in as few parts as possible. If you are going to use a different map that the one provided with this manual, use your own scanner to scan the map as an image file, a resolution of 300 DPI will do. If your map has colors, scan the image in color so that you can later use those colors to separate information from your map into different layers (for ex., forest stands, contour lines, roads...).

DIGITIZATION AND ATTRIBUTION

Digital imaging technologies are replacing the microfilm camera and photocopier as the primary mechanisms for reproducing print and graphic resources. Digitization practices do not necessarily accomplish preservation goals; only a portion of digitization programs in cultural heritage institutions produce preservation-quality results. In 2004, the Association of Research Libraries issued a position paper that supported the creation of preservation-quality digital images, citing the abundance of available standards and best practices. This course concentrates on the state-of-the-art of standards, techniques, metadata, and project requirements for the production of preservation-quality digital images. The course will consider such standards and practices within the larger context of the representation of information through technological remediation.

TOPOLOGY: ERROR DETECTION AND CORRECTION

Errors in power system topology are difficult to handle with classical methods based on state estimation. This paper proposes a pre-processing method for detecting and identifying topology errors and bad measurements before a state estimator solution. Two algorithms are described. These are used at the substation level or at the substation voltage level.

The first algorithm is more suited to systems (substations) where no switch flow measurements are available. Therefore, it does not require any additional measurements other than the ones currently used by the state estimator. The topology error detection uses a bus-branch model. The model relies on the connectivity information and on the impedance values of the non-zero impedance branches.

The second algorithm is more suited to systems (substations) where some switch flow measurements are available. Even if these are not required by the algorithm, using them add redundancy and provides a more reliable solution. The topology error detection uses a bus section—switch device model. The model uses only branch connectivity.

The first algorithm provides better results when no switch flow measurements are available. The second one is superior when some switch flow measurements are available. These algorithms complement each other and they can be use together to cover a much broader scope of the topology error detection and identification problem.

Topology Correction

In healthy subjects, assuming the cerebral cortex is closed off at the brainstem, the boundary of the cortex should be topologically equivalent to a sphere, i.e. it should have no holes or handles. However, because of segmentation errors, surfaces that are generated directly from the binary object produced by the previous step are likely to contain topological holes or handles. These can lead to subsequent problems in producing flat-maps or making 1-1 surface correspondences across subjects. Thus, a graph-based algorithm is applied to force the segmented group of voxels to have spherical topology).

DATA VISUALIZATION, MAP LAYOUT DESIGN AND SYMBOLOGY

Basic ideas about data visualization

When we visualize any kind of geographical information, whether on a computer screen or on a printed map, we are using a visual language to convey that information. The study of signs of a language is called semiology. In the case of a visual language, we have a graphic semiology. This semiology works with the signs of the language that we use to visualize geographical data and helps us understand why and how visual elements serve their purpose of correctly conveying the information from which they are created.

Visual variables

Visual elements have several properties that can be used to transmit information. Depending on the case, some of them might more suitable than others.

These properties are known as visual variables and are applied to the geometric elements used to visualize geographical information. Those elements can be differentiated using the following visual variables.

The use of position is rather restricted in the case of a map, since the real position of the element to be rendered should be respected. It is seldom used.

The shape is defined by the perimeter of the object. This variable is mostly used in the case of point data, using a symbol of a given shape located at the exact coordinates of the point to be rendered. It is difficult to apply to linear symbols and in the case of areal symbols it requires altering the shape of the symbol itself.

Size indicates the dimensions of the symbol. In the case of points, it can be applied by changing the size of the symbol itself. In the case of lines, changing their thicknesses is the most usual way of applying this visual variable on them. It is not used in areal symbols, except in the case of using a texture fill, in which the size variable is applied to the texture and not to the symbol itself.

Size alters how other visual variables are perceived, especially in the case of small sizes.

Texture refers to the pattern used to fill the body of the symbol. It can be applied to lines, using dash patterns, but it is mostly applied to areal symbols.

Color is the most important of all visual variables. Two of its components can be used as individual visual variables themselves: hue and value.

Hue is what we usually call color. That is, the name of the color (blue, red, green, etc.)

Hue can be altered by the hue of surrounding elements, especially in small symbols. Although human perception has a great sensitivity, it might be difficult to identify in small symbols, and it can be wrongly identified if the symbol has other larger ones with different hues in its surroundings.

Value defines the darkness of the color. For instance, light blue and dark blue have the same hue, but they have different value.

Differentiating two symbols by their value can be difficult depending on the type of symbol. It is easier in the case of areal symbols, while in the case of linear and point symbols it depends on their size. Smaller sizes make it more difficult to compare values and to extract the information that the visual variable is trying to convey.

Orientation is applied to point symbols, unless they have some sort of symetry that makes it difficult to identify the orientation of the symbol. For areal symbols, it is applied to their texture. It's not applied in the case of linear symbols.

Maps and cartographic communication

Maps are a method of communication that uses a language with a particular purpose: describing spatial relations. A map is, therefore, a symbolic abstraction of a real-word phenomenon, which implies that it has some degree of simplification and generalization.

The visual language that we have just seen becomes a cartographic language when it is adapted to the particular case of creating maps and knowing its rules is needed to create cartography that is later useful for the map user. All these ideas related to map production form what is known as cartographic design. Cartographic design involves making decisions (in this case, by the GIS user who takes the role of the cartographer). These decisions must be guided by the purpose of the map and the target audience and depending on these factors, the cartographer must decide the projection(which doesn not always have to be the original one of the data), the scale (depending on the level of detail and taking into account the limitations of the data), the type of map, or the symbols to use, among other things.

There are two main types of cartography: base cartography (also called fundamental or topographic) and thematic cartography.

Historically, base cartography represents the classic maps that have been created by cartographers. This type of map serves the purpose of precisely describing *what* is on the surface of the Earth.

Thematic cartography focuses on displaying information about a given phenomenon (a given geographical variable), which can be of any type: physical, social, political, cultural, etc. We exclude from this list those phenomena that are purely topographic, which are the subject matter of base cartography.

We can also say that base cartography represents physical elements (a stream, a coast line, a road, a valley, etc.), while thematic cartography focuses on representing values and attributes.

Thematic cartography uses base cartography (usually included in thematic maps) to help the map user to understand the spatial behavior of the variable being represented, and also to provide a geographical context for it.

Types of information and their visualization

We already know that the thematic component of geographical information can be numeric or alphanumeric and that numeric variables can be nominal, ordinal, intervals, or ratios. Selecting a correct symbology according to the type of information that we are working with is key to producing an effective map. In particular, we must use a visual variable that has the correct properties (levels of organization) for the variable that we want to visualize.

For instance, the associative property and the selective property are of interest just for qualitative information, while size is the only visual variable that we can use that has the quantitative property and therefore, the only one that should be used to represent ratios.

The following are some of the more important ideas about this, referred to the aforementioned types of information.

- Nominal. Nominal information is correctly represented using the visual variable shape. This information shows *what* is found in the different locations of a map, and not *how much* is found, and it is more related to base cartography than to thematic cartography. Using different symbols for point elements and line elements is a common and very effective solution. For the case of areal symbols, hue and texture are the most common solutions.

Alphanumeric information has similar properties, and the same ideas apply to it.

- Ordinal. Since values of the variable define an order, a visual variable with the ordered property is needed to correctly visualize this type of information
- Interval and ratio. Visual variables with the ordered property can be used in this case. However, size is a better choice, as it is the only one which has the quantitative property.

Values are normally grouped into classes so the same value of the visual variable (same size of the symbols or same color value, for instance) is used for different values of the variable that we are visualizing. There are different strategies for this, which try to maximize the information that the map transmits.

The most commons ones are equal intervals, intervals using percentiles or natural intervals (intervals that try to minimize the variance within each class).

MAP THE WORLD IN OPEN SOURCE

QGIS – Formerly Quantum GIS

Highlights: Community All-in-one Cartography Plugins GISGeography Favorite

After the epic GIS software battle in GIS history between ArcGIS vs QGIS, we illustrated with 27 differences why QGIS is undoubtedly the #1 free GIS software package.

QGIS is jam-packed with hidden gems at your fingertips. For example, you can automate map production, process geospatial data, and generate drool-worthy cartographic figures.

There's no other free mapping software on this list that lets you map like a rock star than QGIS.

QGIS Plugins boost this mapping software into a state of epicness. If the tool doesn't exist, search for a plugin developed by the QGIS community. Volunteer effort is key to its success. The QGIS Stack Exchange support is impressively great.

If you're still searching for free GIS software, you'd be insane not to download the free GIS software QGIS. Here's your beginner's guide to QGIS to get your feet wet.

gVSIG

We illustrate in this gvSIG guide and review why we like it SO much:

gvSIG really outperforms QGIS 2 for 3D. It really is the best 3D visualization available in open source GIS.

The NavTable is agile in that it allows you to see records one-by-one vertically.

The CAD tools are impressive on gvSIG. Thanks to the OpenCAD Tools, you can trace geometries, edit vertices, snap and split lines and polygons.

If you need GIS on your mobile phone, gvSIG Mobile is perfect for field work because of its interface and GPS tools.

Whitebox GAT

Unbelievably, Whitebox GAT has only been around since 2009 because it feels so fine-tuned when you see it in action.

There's a hydrology theme around Whitebox GAT. It actually replaced Terrain Analysis System (TAS) – a tool for hydro-geomorphic applications.

Whitebox GAT is really a full-blown open-access GIS and remote sensing software package.

With no barriers, Whitebox GAT is the swiss-army knife of LiDAR data.

The LiDAR toolbox is a life-saver. For example, LAS to shapefile is an insanely useful tool. But you may need a Java update to go in full throttle though.

The cartographic mapping software tools are primitive compared to QGIS.

But overall Whitebox GAT is solid with over 410 tools to clip, convert, analyze, manage, buffer and extract geospatial information.

I find it amazing this free GIS software almost goes unheard of in the GIS industry.

Get more useful knowledge from the Whitebox GAT Open Source Blog.

SAGA GIS

SAGA GIS (System for Automated Geoscientific Analyses) is one of the classics in the world of free GIS software.

It started out primarily for terrain analysis such as hillshading, watershed extraction and visibility analysis.

Now, SAGA GIS is a powerhouse because it delivers a fast growing set of geoscientific methods to the geoscientific community.

Enable multiple windows to lay out all your analysis (map, histograms, scatter plots, attributes, etc). It provides both a user-friendly GUI and API. It's not particularly useful in cartography but it's a lifesaver in terrain analysis.

Closing gaps in raster data sets is easy. The morphometry tools are unique including the SAGA topographic wetness index and topographic position classification. If you have a DEM, and don't know what to do with it – you NEED to look at SAGA GIS.

Overall, it's quick, reliable and accurate. Consider SAGA GIS a prime choice for environmental modeling and other applications.

GRASS GIS

GRASS GIS (Geographic Resources Analysis Support System) was developed by the US Army Corps of Engineers as a tool for land management and environmental planning.

It has evolved into a free GIS software option for different areas of study.

Academia, environment consultants and government agencies (NASA, NOAA, USDA and USGS) use GRASS GIS because of its intuitive GUI and its reliability.

It has over 350 rock-solid vector and raster manipulation tools.

Not awfully useful in cartographic design, GRASS GIS excels primarily as a free GIS software option for analysis, image processing, digital terrain manipulation and statistics.

MapWindow

In 2000, MapWindow was proprietary GIS software. However, it has been made open through a contract with the US EPA called "Basins". At this point, The source code was released to the public.

Now that MapWindow 5 has been released, it surprisingly has some serious punch. For example, MapWindow does about 90% of what GIS users need – map viewer, identify features, processing tools and print layout.

It has some higher level tools such as TauDEM for automatic watershed delineation. While HydroDesktop for data discovery, download, visualization and editing, DotSpatial for GIS programmers. In addition, it has an extensible plugin architecture for customization.

ILWIS

Free GIS software users rejoice. Once commercial GIS software, now turned into open source GIS. ILWIS (Integrated Land and Water Information Management) is an oldie but a goodie.

The extinction-proof ILWIS is free GIS software for planners, biologists, water managers and geospatial users. ILWIS is good at the basics – digitizing, editing, displaying geographic data. Further to this, it's also used for remote sensing with tools for image classification, enhancements and spectral band manipulation.

Over time, it has improved support for time series, 3 analysis and animation. Overall, I found it difficult to do some of the basics like adding layers. However, the documentation is thorough with a pretty decent following for usage.

GeoDa

GeoDa is a free GIS software program primarily used to introduce new users into spatial data analysis. It's main functionality is data exploration in statistics.

One of the nicest things about it is how it comes with sample data for you to give a test-drive. From simple box-plots all the way to regression statistics, GeoDa has complete arsenal of statistics to do nearly anything spatially.

It's user base is strong. For example, Harvard, MIT and Cornell universities have embraced this free GIS software to serve as a gentle introduction to spatial analysis for non-GIS users. From economic development to health and real estate, it's been used as an exciting analytical in labs as well.

uDig

uDIG is an acronym to help get a better understanding what this Free GIS software is all about.

D stands for desktop (Windows, Mac or Linux). You can run uDIG on a Mac.

I stand for internet oriented consuming standard (WMS, WFS or WPS)

G stands for GIS-ready for complex analytical capabilities.

When you start digging into uDig, it's a nice open source GIS software option for basic mapping. uDig's Mapnik lets you import basemaps with the same tune as ArcGIS

Specifically, it's easy-to-use, the catalog, symbology and Mac OS functionality are some of the strong points. But it has limited tools and the bugs bog it down to really utilize it as a truly complete free GIS software package.

OpenJump

Formerly JUMP GIS, OpenJump GIS (JAVA Unified Mapping Platform) started as a first class conflation project. It succeeded. But eventually grew into something much bigger. Because of how its large community effort grew, OpenJUMP into a more complete free GIS software package.

One of its strengths is how it handles large data sets well. Rendering is above-grade with a whole slew of mapping options. For example, you can generate pie charts, plotting and choropleth maps.

OpenJUMP GIS Plugins enhance its capabilities. There are plugins for editing, raster, printing, web-processing, spatial analysis, GPS and databases. Conflating data is another option with a whole lot more from its plugins.

Diva GIS

Biologists using GIS unite! This one specializes in mapping biological richness and diversity distribution including DNA data.

Diva GIS is another free GIS software package for mapping and analyzing data. Diva GIS also delivers useful, every day free GIS data for your mapping needs.

It's possible to extract climate data for all locations on the land. From here, there are statistical analysis and modeling techniques to work with. For the biologist in you, it's worth a long look for biologists around the world. Otherwise, you should be looking at one of the top options above.

FalconView

The initial purpose of FalconView is to be a free and open source GIS software.

Georgia Tech built this open software for displaying various types of maps and geographically referenced overlays.

Now, most of FalconView's users are from the US Department of Defense and other National Geospatial Intelligence Agencies. This is because it can be used for combat flight planning.

In SkyView mode, you can fly-through even using MXD files. It supports various types of display like elevation, satellite, LiDAR, KMZ and MrSID.

OrbisGIS

OrbisGIS is a work-in-progress. Its goal is to be a cross-platform open source GIS software package designed by and for research.

It provides some GIS techniques to manage and share spatial data. OrbisGIS is able to process vector and raster data models.

It can execute processes like noise maps or hydrology process without any add-ons. Orbis GIS Plug-ins are available but are very limited for the time-being.

The developers are still working on the documentation. You may want to look elsewhere until this project gets sturdy up on its feet.

Open Source GIS and Freeware GIS Applications

An open source application by definition is software that you can freely access and modify the source code for. Open source projects typically are worked on by a community of volunteer programmers. Open source GIS programs are based on different base programming languages. Three main groups of open source GIS (outside of web GIS) in terms of programming languages are: "C" languages, Java, and .NET.

The first group would be the group that uses "C" language for its implementation. This is the more mature of the groups of open source GIS, probably for the simple reason that is the group that has been working on GIS software applications the longest and has a long history of resuse of code. The libraries in the "C" group, from the base infrastructure, and include some capabilities like coordinate reprojection that make them very useful and popular. Popular "C" based open source GIS software applications include GRASS, a project started in 1982 by the US Army but is now open source, and QGIS (otherwise known as Quantum GIS).

The second group of Open Source GIS would be the ones that use JAVA as the implementation language. JTS, central library for the Java GIS development, offers some geospatial functions that allow to compare objects and return a boolean true/false result indicating the existence (or absence) of any questioned spatial relationship. Other operators, like Union or Buffer, which are very hard to code, are offered in this group making it very appreciated by GIS developers. GeoTools, Geoserve, and OpenMap, are among the most popular open source GIS in this group of JAVA tools.

The third most influential group of Open Source GIS would be the one that integrates applications that use ".NET" as the implementation language. SharpMap and WorldWind are the most popular of these applications.

Outside of the three major language groups, open source web mapping is another group. Population open source web mapping includes OpenLayers and MapBuilder, widely used due to their simplicity and accessibility.

5

Introduction to Remote Sensing

Remote sensing is the small-or large-scale acquisition of information of an object or phenomenon, by the use of either recording or real-time sensing device(s) that are wireless, or not in physical or intimate contact with the object (such as by way of aircraft, spacecraft, satellite, buoy, or ship). In practice, remote sensing is the stand-off collection through the use of a variety of devices for gathering information on a given object or area. Thus, Earth observation or weather satellite collection platforms, ocean and atmospheric observing weather buoy platforms, the monitoring of a parolee via an ultrasound identification system, Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), X-radiation (X-RAY) and space probes are all examples of remote sensing. In modern usage, the term generally refers to the use of imaging sensor technologies including: instruments found in aircraft and spacecraft as well as those used in electrophysiology, and is distinct from other imaging-related fields such as medical imaging.

Applications of Remote Sensing Data

- Conventional radar is mostly associated with aerial traffic control, early warning, and certain large scale meteorological data. Doppler radar is used by local law enforcements' monitoring of speed limits and in enhanced meteorological collection such as wind speed and direction within weather systems. Other types of active collection includes plasmas in the ionosphere). Interferometric synthetic aperture radar is used to produce precise digital elevation models of large scale terrain.

- Laser and radar altimeters on satellites have provided a wide range of data. By measuring the bulges of water caused by gravity, they map features on the seafloor to a resolution of a mile or so. By measuring the height and wave-length of ocean waves, the altimeters measure wind speeds and direction, and surface ocean currents and directions.
- Light detection and ranging (LIDAR) is well known in examples of weapon ranging, laser illuminated homing of projectiles. LIDAR is used to detect and measure the concentration of various chemicals in the atmosphere, while airborne LIDAR can be used to measure heights of objects and features on the ground more accurately than with radar technology. Vegetation remote sensing is a principal application of LIDAR.
- Radiometers and photometers are the most common instrument in use, collecting reflected and emitted radiation in a wide range of frequencies. The most common are visible and infrared sensors, followed by microwave, gamma ray and rarely, ultraviolet. They may also be used to detect the emission spectra of various chemicals, providing data on chemical concentrations in the atmosphere.
- Stereographic pairs of aerial photographs have often been used to make topographic maps by imagery and terrain analysts in trafficability and highway departments for potential routes.
- Simultaneous multi-spectral platforms such as Landsat have been in use since the 70's. These thematic mappers take images in multiple wavelengths of electro-magnetic radiation (multi-spectral) and are usually found on Earth

observation satellites, including (for example) the Landsat program or the IKONOS satellite. Maps of land cover and land use from thematic mapping can be used to prospect for minerals, detect or monitor land usage, deforestation, and examine the health of indigenous plants and crops, including entire farming regions or forests.

- Within the scope of the combat against desertification, remote sensing allows to follow-up and monitor risk areas in the long term, to determine desertification factors, to support decision-makers in defining relevant measures of environmental management, and to assess their impacts.

Geodetic

- Overhead geodetic collection was first used in aerial submarine detection and gravitational data used in military maps. This data revealed minute perturbations in the Earth's gravitational field (geodesy) that may be used to determine changes in the mass distribution of the Earth, which in turn may be used for geological or hydrological studies.

Acoustic and Near-acoustic

- Sonar: *passive sonar*, listening for the sound made by another object (a vessel, a whale etc); *active sonar*, emitting pulses of sounds and listening for echoes, used for detecting, ranging and measurements of underwater objects and terrain.
- Seismograms taken at different locations can locate and measure earthquakes (after they occur) by comparing the relative intensity and precise timing.

To coordinate a series of large-scale observations, most sensing systems depend on the following: platform location, what time it is, and the rotation and orientation of the sensor. High-end instruments now often use positional information from satellite navigation systems.

The rotation and orientation is often provided within a degree or two with electronic compasses. Compasses can measure not just azimuth (i.e. degrees to magnetic north), but also altitude (degrees above the horizon), since the magnetic field curves into the Earth at different angles at different latitudes.

More exact orientations require gyroscopic-aided orientation, periodically realigned by different methods including navigation from stars or known benchmarks.

Resolution impacts collection and is best explained with the following relationship: less resolution=less detail & larger coverage, More resolution=more detail, less coverage. The skilled management of collection results in cost-effective collection and avoid situations such as the use of multiple high resolution data which tends to clog transmission and storage infrastructure.

PRINCIPLES OF REMOTE SENSING

We perceive the surrounding world through our five senses. Some senses (touch and taste) require contact of our sensing organs with the objects. However, we acquire much information about our surrounding through the senses of sight and hearing which do not require close contact between the sensing organs and the external objects. In another word, we are performing Remote Sensing all the time.

Generally, Remote sensing refers to the activities of recording/ observing/perceiving (sensing) objects or events at far away (remote) places. In remote sensing, the sensors are not in direct contact with the objects or events being observed. The information needs a physical carrier to travel from the objects/events to the sensors through an intervening medium. The electromagnetic radiation is normally used as an information carrier in remote sensing. The output of a remote sensing system is usually an image representing the scene being observed. A further step of image analysis and interpretation is required in order to extract useful information from the image. The human visual system is an example of a remote sensing system in this general sense.

In a more restricted sense, remote sensing usually refers to the technology of acquiring information about the earth's surface (land

and ocean) and atmosphere using sensors onboard airborne (aircraft, balloons) or spaceborne (satellites, space shuttles) platforms.

DATA PROCESSING

Generally speaking, remote sensing works on the principle of the *inverse problem*. While the object or phenomenon of interest (the state) may not be directly measured, there exists some other variable that can be detected and measured (the observation), which may be related to the object of interest through the use of a data-derived computer model. The common analogy given to describe this is trying to determine the type of animal from its footprints.

For example, while it is impossible to directly measure temperatures in the upper atmosphere, it is possible to measure the spectral emissions from a known chemical species (such as carbon dioxide) in that region. The frequency of the emission may then be related to the temperature in that region via various thermodynamic relations. The quality of remote sensing data consists of its spatial, spectral, radiometric and temporal resolutions.

Spatial Resolution

The size of a pixel that is recorded in a raster image – typically pixels may correspond to square areas ranging in side length from 1 to 1,000 metres (3.3 to 3,300 ft).

Spectral Resolution

The wavelength width of the different frequency bands recorded – usually, this is related to the number of frequency bands recorded by the platform. Current Landsat collection is that of seven bands, including several in the infra-red spectrum, ranging from a spectral resolution of 0.07 to 2.1 μm. The Hyperion sensor on Earth Observing-1 resolves 220 bands from 0.4 to 2.5 ìm, with a spectral resolution of 0.10 to 0.11 ìm per band.

Radiometric Resolution

The number of different intensities of radiation the sensor is able to distinguish. Typically, this ranges from 8 to 14 bits, corresponding to 256 levels of the gray scale and up to 16,384 intensities or "shades" of colour, in each band. It also depends on the instrument noise.

Temporal Resolution

The frequency of flyovers by the satellite or plane, and is only relevant in time-series studies or those requiring an averaged or mosaic image as in deforesting monitoring. This was first used by the intelligence community where repeated coverage revealed changes in infrastructure, the deployment of units or the modification/introduction of equipment. Cloud cover over a given area or object makes it necessary to repeat the collection of said location.

In order to create sensor-based maps, most remote sensing systems expect to extrapolate sensor data in relation to a reference point including distances between known points on the ground.

This depends on the type of sensor used. For example, in conventional photographs, distances are accurate in the center of the image, with the distortion of measurements increasing the farther you get from the center. Another factor is that of the platen against which the film is pressed can cause severe errors when photographs are used to measure ground distances. The step in which this problem is resolved is called georeferencing, and involves computer-aided matching up of points in the image (typically 30 or more points per image) which is extrapolated with the use of an established benchmark, "warping" the image to produce accurate spatial data. As of the early 1990s, most satellite images are sold fully georeferenced.

In addition, images may need to be radiometrically and atmospherically corrected. Radiometric correction gives a scale to the pixel values, e.g. the monochromatic scale of 0 to 255 will be converted to actual radiance values.

Atmospheric correction eliminates atmospheric haze by rescaling each frequency band so that its minimum value (usually realised in water bodies) corresponds to a pixel value of 0. The digitizing of data also make possible to manipulate the data by changing gray-scale values.

Interpretation is the critical process of making sense of the data. The first application was that of aerial photographic collection which used the following process; spatial measurement through the use of a light table in both conventional single or stereographic coverage, added skills such as the use of photogrammetry, the use of photomosaics, repeat coverage, Making use of objects' known dimensions in order to detect modifications.

Image Analysis is the recently developed automated computer-aided application which is in increasing use. Object-Based Image Analysis (OBIA) is a sub-discipline of GIScience devoted to partitioning remote sensing (RS) imagery into meaningful image-objects, and assessing their characteristics through spatial, spectral and temporal scale.

Old data from remote sensing is often valuable because it may provide the only long-term data for a large extent of geography. At the same time, the data is often complex to interpret, and bulky to store.

Modern systems tend to store the data digitally, often with lossless compression. The difficulty with this approach is that the data is fragile, the format may be archaic, and the data may be easy to falsify.

One of the best systems for archiving data series is as computer-generated machine-readable ultrafiche, usually in typefonts such as OCR-B, or as digitized half-tone images.

Ultrafiches survive well in standard libraries, with lifetimes of several centuries. They can be created, copied, filed and retrieved by automated systems. They are about as compact as archival magnetic media, and yet can be read by human beings with minimal, standardized equipment.

Data Processing Levels

To facilitate the discussion of data processing in practice, several processing "levels" were first defined in 1986 by NASA as part of its Earth Observing System and steadily adopted since then, both internally at NASA and elsewhere; these definitions are:

Level	*Description*
0	Reconstructed, unprocessed instrument and payload data at full resolution, with any and all communications artifacts (e.g., synchronization frames, communications headers, duplicate data) removed.
1a	Reconstructed, unprocessed instrument data at full resolution, time-referenced, and annotated with ancillary information, including radiometric and geometric calibration coefficients and georeferencing parameters (e.g., platform ephemeris) computed and appended but not applied to the Level 0 data (or if applied, in a manner that level 0 is fully recoverable from level 1a data).
1b	Level 1a data that have been processed to sensor units (e.g., radar backscatter cross section, brightness temperature, etc.); not all instruments have Level 1b data; level 0 data is not recoverable from level 1b data.
2	Derived geophysical variables (e.g., ocean wave height, soil moisture, ice concentration) at the same resolution and location as Level 1 source data.
3	Variables mapped on uniform spacetime grid scales, usually with some completeness and consistency (e.g., missing points interpolated, complete regions mosaicked together from multiple orbits, etc).
4	Model output or results from analyses of lower level data (i.e., variables that were not measured by the instruments but instead are derived from these measurements).

A Level 1 data record is the most fundamental (i.e., highest reversible level) data record that has significant scientific utility, and is the foundation upon which all subsequent data sets are produced. Level 2 is the first level that is directly usable for most scientific applications; its value is much greater than the lower levels. Level 2 data sets tend to be less voluminous than Level 1 data because they have been reduced temporally, spatially, or spectrally. Level 3 data sets are generally smaller than lower level data sets and thus can be dealt with without incurring a great deal of data handling overhead. These data tend to be generally more useful for many applications. The regular spatial and temporal organization of Level 3 datasets makes it feasible to readily combine data from different sources.

URBAN TEXTURAL ANALYSIS FROM REMOTE SENSOR DATA

Despite the new generation of very high spatial resolution sensor data (IKONOS from 1999 and QuickBird from 2001), predicted improvements in classification accuracy of urban land covers (and subsequent inference of urban land use) have yet to materialize substantially. Much of the obstruction to quality information extraction is still due to the traditional limitations of classifying image data representing urban areas: the high spatial arrangement of complex urban features and how to configure multispectral responses from land cover features into organized urban land-use categories (Barr, Barnsley, and Steel 2004). When launched, the desired objective of high spatial resolution sensor data was for increased clarity of terrestrial features, especially urban objects, by reducing per-pixel spectral heterogeneity and thereby improving land cover identification.

Clarity is certainly more evident in these finer-scale data than those from preceding sensors, but paradoxically this greater level of detail is also translated into many more unique per-pixel spectral combinations. For example, the residential land-use category can now be defined from much wider spectral variations, representing minute compositional mixtures of urban land covers, such as roads, houses, grasses, trees, bare soil, shrubs, and swimming pools, each conceivably a different residential land-use category. Following on, another limitation for improved information extraction from high spatial resolution sensor data is the reliance on techniques using traditional per-pixel spectral differentiation. To us this seems counterintuitive and we would like to see more neighbourhood-related methods, using textural and spatial parameters when dealing with fine-resolution image data.

Where traditional spectral approaches are designed to identify homogeneous features regardless of shape, textural and spatial algorithms measure both the variance within and the geometric configuration of whole urban objects, respectively. As a contribution to the growing literature, we outline an object-based pattern recognition technique that accommodates the concept of lacunarity for characterizing the textural properties of urban land cover (and therefore inferring land use) from high spatial resolution image data. In doing so, we consolidate the utility of geometric models not only for image data but for all discrete and textural spatial representations (Zhao and Stough 2005). Indeed, the ability to characterize the shapes of individual and groups of objects is a rapid area of research in computational geometry and at the heart of the recent developments in object-based models in many geographic information system algorithms.

Recall that remote sensor data are composed of multispectral pixel vectors that represent geographical objects and their relative configuration. We strongly adhere to the paradigm that geometric patterns, such as lacunarity, are valuable precursors for functional processes; in our application, the texture and spatial orientation of land-cover patterns derived from remote sensor data are both forerunners for analysing land-use juxtaposition and dynamic urban processes.

Lacunarity Approach

The lacunarity of an object is the counterpart of its fractal dimension. Lacunarity methods for urban analysis, and indeed many other applications in geospatial research, have already been reported by a number of researchers.

Essentially, lacunarity is related to the spatial distribution of gap or hole sizes. For low-lacunarity measurements, all gap sizes are the same and geometric objects are deemed homogeneous; conversely, for high-lacunarity gap sizes are variable and objects are therefore heterogeneous.

In other words, the variance or texture of gap sizes within the spatial delineation of geometric objects determines the level of lacunarity. Of course, textures that are homogeneous at small scales can be quite heterogeneous at large scales, and vice versa; therefore, lacunarity can be considered a scale-dependent measure of texture. Methods for calculating the lacunarity of objects were first given, in general terms, by Mandelbrot (1983) and were later implemented by various computer algorithms.

Work by Myint and Lam (2005b) developed two modified lacunarity algorithms: the binary approach, which was first introduced by Plotnick, Gardner, and O'Neill (1993), and a gray-scale routine, initially devised by Voss (1986) and used to test the effectiveness of lacunarity on high spatial resolution sensor data. This same desire to extract urban objects from fine-scale sensor data also forms the basis of this study, where we examine modifications to a differential box counting algorithm, first formulated by Dong (2000b).

Our study will also introduce two different gliding box approaches. The first uses overlapping boxes, in which the gliding box moves to a pixel next to the previous position, and the second uses skipping boxes, in which the gliding box skips the entire coverage of the previous box before moving to the next position. As a background, and according to the gliding box algorithm proposed by Allain and Cloitre (1991), n(M,r) can be defined as the number of gliding boxes with radius rand mass M. The computations of lacunarity values are given by worked examples where the overlapping box method is demonstrated by a 4 x 4 image or local window, while a 6 x 6 image is used to illustrate the skipping box method. The 3 x 3 gliding box used in both is the base of the cube box and is always an odd number to allow the computed value to be assigned to a central cell.

A column with more than one cube box may be required to cover the maximum image intensity values by stacking cube boxes on top of each other.

The number of cube boxes required to cover the image intensity surface depends on the pixel values in the 3 x 3 gliding box. Intensity values using the example are calculated at the first, second, third, and fourth positions of the cube boxes (overlapping and skipping boxes).

The minimum and maximum pixel values are 7 and 18, respectively, at the first position of the gliding box. With a cube box of 3 x 3 x 3, these values fall in box number 3 (value of u) and 6 (value of v), respectively. The relative height of the column is then 6-3 + 1 = 4 (u-v + 1). In the same way, we can compute the required parameters for all positions of the cube boxes as follows: For the second position of the gliding box, u= 7, v= 1 (the relative height of the column is (u-v + 1) or 7-1 + 1 = 7). For the third position, u=6; v=5 (the relative height is 6-5 + 1 =2).

RESEARCH DESIGN

Data and Study Area

We applied our technique to an IKONOS sensor image data of 4m spatial resolution across all four of its channels: blue (0.45-0.52 [micro]m), green (0.52-0.60 [micro]m), red (0.63-0.69 [micro]m), and near infrared (0.76-0.90 [micro]m). The image represents the settlement of Norman, Oklahoma and was captured on March 20, 2000. Only a subset of this IKONOS sensor image (614 x 447 pixels) covering a central portion of the metropolitan area was used to identify three land cover/land-use classes: grassland, commercial, and residential—all three capable of being delineated using manual interpretation.

Local Window

At the onset, it is important to note that the characteristic scale (Lark 1996) is an important parameter to be considered for effective identification of land covers with different texture appearances. A characteristic scale is the minimum distance between two pixels that completely covers a texture. In our study, anything between 21 pixels (84 m) and 27 pixels (108 m) was considered large enough to cover textures that represent all of the classes, especially complex residential and commercial land uses derived from the IKONOS multispectral image. Hence, 9 x 9, 15 x 15, 21 x 21, 27 x 27, 33 x 33, and 39 x 39 local window sizes were used to determine the optimum scale with which to identify land-cover classes.

From a previous study (Myint and Lam 2005b) that was based on a lacunarity approach designed by Voss (1986), we demonstrated that smaller gliding boxes have more discriminatory power than larger gliding boxes. And as a result, a small 3 x 3 x 3 gliding box is used in this study and is the very basis for the chosen window sizes dimensionalized by a factor of 3. Hypothetically, a window size should be small enough to cover only single land-cover features but large enough to guarantee sufficient spatial/textural information for the characterization of land-cover types.

In that case, a mirror extension of (w-1)/2 pixels around the image is necessary before beginning computation. What happens is that the algorithm is designed to extend automatically the image with (w-1)/2 pixels all around if the selected window size is w. As such, the size of an extended image is the original image size + (window size-1). Mirror extension is designed to copy the second-last row or column and add the next-to-last row and column, respectively. Then copy the third-last row and column and add the next-to-the-second-last row and column, respectively, and so on depending on the number of rows and columns required for the extension.

This mirror extension procedure is demonstrated by a hypothetical image and is considered more effective than other widely used methods, including those adding zero, one, last row/

column data, or the mean values in the extended areas. Our favoured approach is also more accurate than other interpolation methods (e.g., kriging, inverse distance) because it does not alter any values in the original image. By using the mirror extension, we obtain the same mean and standard deviation statistics within the extended area.

Training Samples

With knowledge of the local area and ground checks, training samples were carefully selected to represent the three land-cover/ land-use classes: grassland, commercial, and residential— descriptive statistics of their brightness values are shown. It is generally suggested that the standard deviation of training data should be around, or even less than, 10% of the mean coefficient of variation to guarantee homogeneous distributions vital to traditional spectral-based classifiers. Compared with the standard deviation, which is expressed in absolute terms, the coefficient of variation is considered a more appropriate measure for the comparison of data distributions across variables.

The standard deviation of the commercial and residential land uses is 37.1% and 34.3% of their respective mean values. At this stage, it is important to remember that groups of adjacent pixels generally exhibit positive autocorrelation, or at least display a high probability of having similar digital numbers. In our calculations, the grassland (Moran's l = 0.49; Geary's C = 0.43) sample exhibited a positive autocorrelation, whereas the commercial (Moran's l = 0.16; Geary's C=0.83) and residential (Moran's l = 0.09; Geary's C = 0.91) samples displayed random spatial arrangements with no patterns of clustering or homogeneity.

Band Combinations and Classifications

Finally, lacunarity algorithms, based on both skipping and overlapping boxes, were applied to transform bands 2, 3, and 4 of the IKONOS image. In addition, a number of layer stacks involving both lacunarity-transformed bands and original bands were generated in the following combinations:

- three original bands,
- three lacunarity-transformed bands,
- combination of three original bands and three lacunarity-transformed bands,
- combination of three original and one lacunarity-transformed band, and
- combination of two original bands and two lacunarity-transformed bands.

All window sizes were applied to stacks of three original bands and three lacunarity-transformed bands, as well as the combination of three original bands and one lacunarity-transformed band. Regarding the three separate lacunarity-transformed bands, only the 27 x 27 window was used, primarily to observe whether satisfactory accuracy levels were attained without the original bands. In the case of the combination of the three original bands and one lacunarity-transformed band, the lacunarity-transformed band 3 was used to test whether one lacunarity-transformed band could improve the accuracy of the three original bands.

Experiments were conducted with all window sizes using lacunarity-transformed band 3, as well as testing lacunarity-transformed band 4 with the original bands using the 27 x 27 window for significant differences between lacunarity-transformed band 3 and lacunarity-transformed band 4. In the case of the combination of two original bands and two lacunarity-transformed bands, band 3 and band 4 were tested using the 27 x 27 window. All code for the modified lacunarity approaches was written in the C + + programming language.

Reference Map for Error Matrix

When examining the effectiveness of different classification algorithms, "wall-to-wall" comparisons were used by checking every pixel in the image. The selected area covered 274,458 pixels (447 x 614) representing all three urban classes: grassland, commercial, and residential. Reference classes were collected by manual interpretation, using sound local area knowledge and a thorough ground survey, and then digitally delineated with negligible positional error. Overall accuracies produced by the combination of three original bands and three lacunarity-transformed bands using the overlapping box algorithm based on the 9 x 9, 15 x 15, 21 x 21, 27 x 27, 33 x 33, and 39 x 39 window sizes were 59.36%, 61.72%, 61.86%, 61.01%, 59.94%, and 61.65%, respectively. They demonstrate that overall accuracy increases slightly with expanding window sizes, despite a slight decline when using the 33 x 33 window, and it is fairly reasonable to suggest that the 15 x 15 and 21 x 21 window sizes are effective for this analysis.

However, overall accuracies produced by the combination of three original bands and three lacunarity-transformed bands using the skipping box algorithm based on the 9 x 9, 15 x 15, 21 x 21, 27 x 27, 33 x 33, and 39 x 39 window sizes were 58.31%, 59.29%, 57.40%, 56.41%, 56.92%, and 58.31%, respectively. Different window sizes produced inconsistent overall accuracies, and hence it is difficult to determine the optimal window size. In comparison, the overlapping box algorithm outperforms the skipping box algorithm, and because the overall accuracy of the original bands is 58.38%, it can be concluded that lacunarity using the overlapping box algorithm improved the classification accuracy of the original bands. However, we feel that accuracy was not significantly improved.

Overall accuracies produced by the combination of three original bands and one lacunarity-transformed band (band 3) using the overlapping box algorithm based on the 9 x 9, 15 x 15, 21 x 21, 27 x 27, 33 x 33, and 39 x 39 window sizes were 58.49%, 60.74%, 60.94%, 60.43%, 60.29%, and 61.35%, respectively. From this, we can deduce that there is only a slight difference between adding the one lacunarity-transformed band (band 3) to the original bands and adding three lacunarity-transformed bands to the original bands. However, it was also found that the combination of three original bands and three lacunarity-transformed bands produced higher accuracies.

Overall accuracies generated by the combination of three original bands and one lacunarity-transformed band (band 4) using the overlapping box algorithm for the 15 x 15 and 27 x 27 window sizes (58.14% and 58.99%, respectively) were lower than the combination with lacunarity-transformed band 3. Overall accuracies produced by the combination of three original bands and one lacunarity-transformed band (band 3) for the skipping box algorithm using the 9 x 9, 15 x 15, 21 x 21, 27 x 27, 33 x 33, and 39 x 39 window sizes were 58.87%, 60.07%, 58.60%, 57.37%, 56.93%, and 57.69%, respectively. The results show that there is no difference between adding three lacunarity-transformed bands to the original bands or adding one lacunarity-transformed band (band 3) to the original bands.

This may be mainly because the skipping box algorithm is not particularly effective in identifying texture features from image data. As mentioned earlier, we also tested the lacunarity-transformed bands generated by the overlapping box approach, individually.

These texture bands on their own gave slightly lower accuracy (57.80%) than the individual original bands, which tends to imply that spatial information alone may not be effective in identifying land-use/land-cover classes (an area we intend to research further). We also examined the combination of two original bands (bands 3 and 4) and two lacunarity-transformed bands (again bands 3 and 4) using the overlapping box algorithm based on the 27 x 27 local window. The results from this approach produced an accuracy level as high as the combination of three original bands and three lacunarity-transformed bands. This is probably because IKONOS bands 2 and 3 are visible bands and therefore highly correlated.

Urban remote sensing is currently experiencing a paradigm shift away from spectral-only classification and toward the identification of urban objects using spatial metrics and neighbourhood-based pattern recognition. Most of the conceptual restructuring is underscored by the availability of very high spatial resolution imagery, but also by the practical necessity of generating base maps of urban land use. Our work reported in this chapter is a contribution to research on spatial and textural identification of urban objects. We explored the utility of lacunarity to measure urban heterogeneity where results suggest that the overlapping box approach is more effective than the skipping box alternative.

We documented that there is no significant difference between window sizes, except for the case of the 9 x 9 window, and it might be reasonable to conclude that the 15 x 15, 21 x 21, and 27 x 27 windows are the most effective sizes in our study. It was also found that the combination of two original bands (bands 3 and 4) and two lacunarity-transformed bands (bands 3 and 4) is as accurate as the combination of three original bands and three lacunarity-transformed bands.

The combination of three original bands and three lacunarity-transformed bands is only a slight improvement on the combination of three original bands and one lacunarity-transformed band (band 3). However, what is certain is that the original bands alone or the purely lacunarity-transformed bands are not effective for this type of land-use and land-cover mapping.

Although variable, our results are very much in line with other documented work on textural and spatial characterization of high spatial resolution sensor data of urban land cover. It seems metrics, be they indices of dispersion, contagion, fractal, or lacunarity, are highly sensitive to the initial image segmentation, and inextricably affected by site, time, and scene.

Another reason for small and variable improvements in urban representation is the conceptual gulf between the computational limitations of the discretized remote sensor data model and the heterogeneous and dynamic nature of the urban landscape.

Such a gulf may prove difficult to bridge immediately with either spectral or textural/spatial indices, and a more integrative approach may ultimately be more appropriate. However, for now our work on lacunarity demonstrates another contribution to the potential for spatial metrics to characterize the increasingly finer spatial resolution of remote sensor data. We hold firm to our belief that highly detailed, spatially heterogeneous urban land cover can be measured by algorithms that are equally sensitive to geometric fluctuations.

APPLICABLE TO REMOTE SENSING

The Daubert standards reviewed above will only be applied to remotely-sensed data presented through expert testimony. Remote sensing evidence will be subject to several FRE, which are applicable whether or not an expert is called to testify.

Relevancy, Authentication, and Foundation Any evidence, scientific or otherwise, must be found relevant to the case, meaning that it must make a consequential fact more or less probable than would be deemed otherwise. If used to aid witness testimony, the map must help the trier of fact understand the testimony. Once evidence is found to be relevant, it must be authenticated. Extrinsic authentication is necessary unless the map fulfills one of the self-authentication exceptions listed in Rule 902 of the FRE. A map published by the government, for instance, is self-authenticating under Rule 902(5).

Finally, the evidence must have an adequate foundation; it must be accurate and reliable. If accuracy cannot be confirmed, courts will not admit the evidence. Of these provisions, the main evidentiary hurdle for digital maps is reliability. Courts will ask where the information in the map originated, how the information was transformed into digital form, and how the map itself was created. Since computers create digital maps, the maps will face reliability challenges as computer evidence. Courts, for instance, will inquire into "computer programming errors, equipment malfunction, data entry errors, and the volume of electronic data."

Courts will also closely consider the authenticity of digital maps, particularly where the map does not meet one of the aforementioned self-authentication exceptions. As such, courts will follow Rule 901(a), requiring proof that the evidence is what its proponent claims it to be. According to Rule 901(b)(9), parties must prove that evidence encompassing a process or system, such as maps depicting remotely-sensed data, must produce an accurate result. To satisfy these rules, the experts who collected the remotely-sensed data should describe how the process operates and their involvement. Experts should also reference the data to ground information ('ground-truthing'), aerial photographs, and other maps. Logs and records of the progression from collection to presentation of the data would also verify authenticity. Technologies including steganography and cyclic redundant checksum are continually being developed to assist in ensuring the authenticity of digital imagery.

Hearsay Issues

If a map, chart, or other media is admitted to make an assertion, the evidence may be objected to on hearsay grounds. For example, remotely-sensed data could be used to create a map depicting high levels of pollution in a stream adjacent to the defendant's property. If the map is admitted to assert that the defendant caused such pollution, it may meet with a hearsay objection. If the evidence is found to be hearsay, it will only be admissible if it can be categorized as an exception to the hearsay rules. For example, Rules 803(6) and 803(8) will allow the admission of hearsay evidence that was generated by computer for use as a business or public record.

Data Characterization

A final set of rules that may pertain to the use of remotely-sensed data involve the presentation of the evidence in the courtroom. Rule 1006 allows the admission of charts, summaries, and calculations that depict a body of data too voluminous to itself be admitted into evidence for practical reasons. To avoid potential problems with admission under this rule, experts should testify that the data was correctly translated into these summary forms. If the evidence is admitted without the verification of expert testimony, Rule 1002 requires that the underlying data be admissible. For example, if a chart includes data derived from satellite photos, courts or opposing attorneys could bar the admission of the chart if the original photos do not also meet the standards of admissibility.

Constitutional Hurdles

Besides Daubert and the FRE, the Constitution presents another obstacle that remote sensing data must overcome for admission into federal courts. The main constitutional issues facing remote sensing data are allegations of invasions of privacy and warrantless searches. The Fourth Amendment states that "the right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated." Two Supreme Court cases, Dow Chemical Co. v. United States and Kyllo v. United States, address the application of the Constitution to remote sensing data.

In Dow Chemical Co., the Court held that enhanced aerial photographs of an industrial facility taken by the EPA were admissible under the Fourth Amendment. The Court found that though commercial areas receive constitutional privacy protection, this protection does not extend to the outdoor areas of industrial complexes.

The Court also found that homes and their outside areas receive a higher level of protection than commercial areas. Still, in dicta the Court stated, "surveillance of private property by using highly sophisticated surveillance equipment not generally available to the public, such as satellite technology, might be constitutionally proscribed absent a warrant." The Court feared that technology providing information not available to the naked eye would reveal intimate details, for example, imaging that could reveal actions occurring inside a building (e.g., conversations behind closed doors or people transporting documents). Despite this concern, the Court noted that photos enhancing human vision were still admissible, provided that they do not reveal such intimate details.

The Supreme Court's latest decision regarding remote sensing data's privacy and search issues is Kyllo v. United States. Kyllo involved a police officer who used a thermal imaging device to detect heat emissions from a suspect's home. Declaring this search unconstitutional, the Court held that when "the Government uses a device that is not in general public use, to explore details of the home that would previously have been unknowable without physical intrusion, the surveillance" is unconstitutional. As in Dow Chemical Co., the Court emphasized that homes receive a high level of privacy protection under the Constitution. The Court held that, in the home, "all details are intimate details," strongly indicating that any information obtained by remote sensing data from a home's interior without a warrant would be inadmissible.

The Court did not define "general use" technology in either Dow Chemical Co. or Kyllo. Lower courts are left to speculate on what level of use might rise to this standard. For example, remote sensors that track wetland deterioration might be deemed "general use" technology if they are routinely used by the government, or if the public accepted their use. But if the device determined that someone illegally filled in a wetland in his or her backyard, that information could be inadmissible. The main lesson that can clearly be drawn from Dow Chemical Co. and Kyllo is that, in the absence of a warrant, remote sensing data will only gain courtroom admission if it does not include intimate details of commercial activity or any details from private homes.

SIGNATURE IN REMOTE SENSING

The knowledge of spectral signatures is essential for exploiting the potential of remote sensing techniques. This knowledge enables one to identify and classify the objects of agricultural resources. It is also required for interpretation of all remotely sensed data, especially in agricultural resource data whether the interpretation is carried out visually or using digital techniques.

It also helps us in specifying requirements for any remote sensing mission *e.g.* which optimal wave length bands to be used or which type of sensor will be best suited for a particular task (agricultural survey). All objects of agricultural resource on the surface of the earth have characteristic spectral signatures. The average spectral reflectance curves (or) spectral signatures for three typical earth's features; vegetation, soil and water. The spectral reflectance curves for vigorous vegetation manifests the "Peak-and valley" configuration. The valleys in the visible portion of the spectrum are indicative of pigments in plant leaves. Dips in reflectance that can be sent at wavelengths of 0.65 mm, 1.4 mm and 1.9 mm are attributable to absorption of water by leaves. The soil curves show a more regular variation of reflectance.

Factors that evidently affect soil reflectance are moisture content, soil texture, surface roughness and presence of organic matter. The water curves shows that from about 0.5 mm, reduction in reflectance with increasing wavelength, so that in the near infrared range, the reflectance of deep clear water is virtually zero (Mather, 1987) However, the spectral reflectance of water is significantly affected by the presence of dissolved and suspended organic and inorganic material and by the depth of the water body.

Determinations of spectral signatures implies basic under standing of interaction of electromagnetic radiation with agricultural resources objects. This is also necessary for analysing and designing sensor systems for agricultural survey.

Sensor Systems in Remote Sensing

In remote sensing the acquisition of data is depending upon the sensor system used. Various remote sensing platforms (Aircraft, Satellite) are equipped with different sensor systems.

Sensor is a device that receives electromagnetic radiation, converts it into a signal and presents it in a form suitable of obtaining information about the land or earth resource as used by an information gathering system. Sensor can be grouped, either on the basic of energy source. They are as classified.

Active sensor: An active sensor operates by emitting its own energy, which is needed to detect the various phenomena (*e.g.* RADAR, camera with a flash gun).

Passive sensor: The operation of passive sensor is dependent of the existing sources of energy, like sun (*e.g.* photographic systems, multispectral scanners).

The given sensor system of camera are in agricultural survey.

Photographic Cameras

The photographic system, having conventional camera with black and white photography, is the oldest and probably, so far, the most widely used sensor for recording information about ground object.

Photographic cameras have been successfully used in aircraft platform remote sensing. In this system, the information is limited to size and shape, as the films used are sensitive only to visible region of spectrum. The response of black & white films is about 0.4-0.7 mm for infrared imagery, films with response extending up to 0.9 mm are available.

Return Beam Vidicon (RBM)

This is very similar to a television camera. In such a system, a fixed camera lens on a photosensitive semi-transparent sheet forms the ground image. This image is created on the surface as electrical change or potential. The TV cameras are the best example of high resolution, operated in space for resource survey was the RBV used in LAND SAT series. On LAND SAT I, II and III RBV cameras were used, each corresponding to a different wavelength band 0.475-0.585 mm (green), 0.580-0.690 mm (red) and 0.690-0.830 mm (near infrared). The Indian experimental remote sensing satellite, Bhaskara-I and II carried a two-band TV camera system, Multispectral imagery was produced in LAND SAT and Bhaskara by using separate camera tubes of each band and selecting the spectral band with appropriate filters.

Optical-mechanical Scanners

This imaging system has the advantage that any set of desired spectral bands can be selected with appropriate filter and detector combinations. The mostly widely used sensor in this category is the MSS on LAND SAT series. MSS has four spectral bands, covering form 0.5-to 1.1 mm region. MSS operates on the principle of scanning successive lines at right angles to the flight path by means of a rotation or oscillating optical system. The radiation levels along the lines are recorded by appropriate sensor elements. When used in the visible band, the collected light can be split by the optics and separately filtered and recorded, giving simultaneous multispectral recording from the one instrument. MSS can record in any part of ultraviolet to near IR window. They are use also in the thermal IR windows.

Radar and Microwave Sensors

The acquisition of data in microwave region has been possible since 1950s but its application to natural resources is considerably less developed, as compared to the visible and IR image interpretations. Microwave sensors have distinct advantages because they are unaffected by atmospheric conditions and are thus able to penetrated smoke, clouds, haze and snow. Under this system, Plan Position Indicator (PPI), Side Looking Air borne Radar (SLAR) and Synthetic Aperture Radar (SAR) can be grouped. These systems offer day and night as well as all weather capability and ability to penetrate a cover of vegetation.

Advance Remote Sensors

Linear Imaging and Self Scanning Sensors (LISS) are the advanced imaging systems. This type of scanning sensor are used an array of solid-state devices. The array may be made of photo-diodes, phototransistors or Charge-Coupled Devices (CCDs). In the LISS, the optics focuses a strip of terrain in the cross-track into the sensor array. The image from each detector is stored and shifted out sequentially to receive a video signal. The SPOT (Satellite Probatorie d' Observation de la Terra) and IRS (Indian Remote

Sensing Satellite) series carry such solid-state sensor systems, which are also known as push-broom scanners. The IRS IC most advanced satellite, carries an improved sensor system. Besides carrying a sophisticated LISS-III camera, it has a Panchromatic camera (PAN) and a Wide Field Sensor (WiFS). The PAN has been designed to provide data with a spatial resolution of 5.8m in stereo mode, with a ground swath of 70km, whereas WiFS provides data in two spectral bands, with a spectral resolution of 188m and a ground swath of 180km.

DATA ACQUISITION TECHNIQUES IN REMOTE SENSING

Remote sensing is the acquisition of information about an object or phenomenon, without making physical contact with the object. In modern usage, the term generally refers to the use of aerial sensor technologies to detect and classify objects on Earth (both on the surface, and in the atmosphere and oceans) by means of propagated signals (*e.g.* electromagnetic radiation emitted from aircraft or satellites).

There are two main types of remote sensing: passive remote sensing and active remote sensing. Passive sensors detect natural radiation that is emitted or reflected by the object or surrounding area being observed. Reflected sunlight is the most common source of radiation measured by passive sensors. Examples of passive remote sensors include film photography, infrared, charge-coupled devices, and radiometers.

Active collection, on the other hand, emits energy in order to scan objects and areas whereupon a sensor then detects and measures the radiation that is reflected or backscattered from the target.

RADAR is an example of active remote sensing where the time delay between emission and return is measured, establishing the location, height, speed and direction of an object.

Remote sensing makes it possible to collect data on dangerous or inaccessible areas.

Remote sensing applications include monitoring deforestation in areas such as the Amazon Basin, glacial features in Arctic and Antarctic regions, and depth sounding of coastal and ocean depths.

Military collection during the cold war made use of stand-off collection of data about dangerous border areas. Remote sensing also replaces costly and slow data collection on the ground, ensuring in the process that areas or objects are not disturbed.

Orbital platforms collect and transmit data from different parts of the electromagnetic spectrum, which in conjunction with larger scale aerial or ground-based sensing and analysis, provides researchers with enough information to monitor trends such as El Niño and other natural long and short term phenomena. Other uses include different areas of the earth sciences such as natural resource management, agricultural fields such as land usage and conservation, and national security and overhead, ground-based and stand-off collection on border areas.

By satellite, aircraft, spacecraft, buoy, ship, and helicopter images, data is created to analyze and compare things like vegetation rates, erosion, pollution, forestry, weather, and land use. These things can be mapped, imaged, tracked and observed. The process of remote sensing is also helpful for city planning, archaeological investigations, military observation and geomorphological surveying.

Data Acquisition Techniques

The basis for multispectral collection and analysis is that of examined areas or objects that reflect or emit radiation that stand out from surrounding areas.

THE PROMISE OF REMOTE SENSING

Remote sensing advocates argue that the flaws inherent in data sets traditionally used in land cover analysis can be corrected by using satellite images to measure the "real" cover of the ground. These traditional data, which include census data and modelled estimates, are indeed problematic and their unclear genealogies and biases make them analytically questionable. For example, many allegedly authoritative assessments of global forest and grassland cover depend upon data reported by the World Resource Institute, the Global 2000 Report, and World Watch research papers, which are, in turn, based upon national self-reporting and state census (World Resources Institute 1992). These are commonly biased through the interests of the state to reflect rhetorical and propagandistic coverage requirements; where forests are politically useful, for example, they often appear in data sets. These data are also collected under dissimilar methods, which makes comparison difficult. Moreover, these data are assembled in categories that are inherited relics of historical environmental management that reflect arbitrary conveniences in administration. Understanding land cover dynamics is difficult if one is forced to rely on these kinds of statistics.

Efforts in geographic information science and remote sensing, on the other hand, according to its advocates, can "develop methodologies for global land cover classifications that are objective [and] reproducible...". Poor sampling and biased reports will, in this way, be rectified through a controlled analysis of remotely sensed data to create accurate category sets. The resulting images, freed from the bias of politics, will appear to provide "a message without a code" and knowledge of the Earth's surface free from the prejudice of authorship. By ground-truthing, assessing for accuracy, and employing recently developed post-classification corrections, the resulting classification supposedly becomes a good match between the image and the land covers the analyst is looking for on the ground.

Although such remote sensing methods show promise, associated landscape classification faces technical obstacles. Estimates of land cover change based solely on satellite data, for example, are limited in the richness of vertical information, especially in thickly layered and complex ecologies. Similarly, mismatches in spatial and temporal scale and resolution make estimates of change difficult to assess and compare. Even so, remote sensing promises comprehensive inventory and analysis, free of the subjective filters of society and politics.

Classification as Metaphor

But the categories of remote sensing analysis are no freer from inherited meaning than those of state-level aggregated statistics. Rigorous image processing can mathematically cluster pixels with like reflectance values but must ultimately rely on an assignment of meaning to such clusters that is inter-textual; it has been inherited from other information and knowledge sources and is historically arbitrary. The arbitrary character of these land cover categories is rooted in their role as metaphors for complex biotic and non-biotic landscape features, which give common meaning to disparate objects. In landscapes characterized by heterogeneity, features are gathered into or excluded from common types by deciding that some relevant characteristic for distinction makes them essentially the same or different. Such characteristics ultimately must be drawn from the metaphoric vocabulary of the analyst.

The Inevitability of Metaphors in Classification

Consider a typical false colour composite image (assembled from Green, Red, and Infrared bands) of a set of agricultural villages adjacent to a hilly area. This visualization of a complex data set shows high levels of variability in digital pixel values across three spectral ranges of reflectance. Certain features in the image will be apparent to the trained eye; the red hills may be heavily vegetated, the winding gray paths extending from the hills might suggest dry streambeds, and darker blue polygons might be towns. The act of interpretation has already begun before any formal analysis, clustering, or classification is conducted. The targets of our classification, those things we already consider real, are winnowed in advance: forests, streambeds, and towns.

To assign formal qualitative meaning to what we see in the image still requires a vast number of decisions about what pixels are most like one another, but the metaphors for understanding are already in place. Unsupervised classification, for example, uses a clustering algorithm to form clusters from the undifferentiated mass of pixels, in an apparently objective fashion. The algorithm guarantees that, based on scrutiny of a multi-dimensional histogram, pixels share membership in categories of "like" reflectance. The number of clusters must be decided in advance, although in most cases, the standard procedure is to overestimate the number of categories, cluster the data, and then aggregate pixel groups together into meaningfully "different" sets. To do so, the analyst surveys the ground, and perceives that "these pixels represent landscapes that are similar, or enough like something I have seen, or seen defined elsewhere; thus, they are effectively the same." Defined by such standards, pixels are lumped together and aggregated to become an authentic image of the Earth's surface. Multiple unique locations become "forest," for example, in the moment that they are named in the same way. This then, is an act of metaphoric deployment. There is no true forest per se, only multiple discrete things brought together as if they were the same; by conjoining and naming them together, they become the same, however, and are accepted as literal in the moment of mapping.

Supervised classification works in a reverse fashion; places known to be what the analyst recognizes as "forest," for example, are used to create training sites, a cluster of pixels of a known cover type. Reflectance signatures are derived from these sites and undifferentiated pixels are categorized by similarity to the training signature, using a range of mathematical decision rules, such as minimum or maximum values or likelihood based on probability. Here again, the concept or trope of forest, an ideal type, is used to name, constitute, and bound cartographic entities, formed prior to analysis. Iterative techniques, such as Isodata clustering, neural networks, and tree-based classifiers that partition and "prune" categories, combine elements of supervised and unsupervised classification, going back and forth between composited images and raw data to produce coherent clusters. Like unsupervised classification, these clusters must still be aggregated and named.

Like supervised classification, these methods utilize carefully chosen training sites—areas of already known land cover—to check against resulting classifications.

Modeling Metaphors: Fuzzy Sets and Land Characteristics Databases

Methods do exist that take seriously the ambiguous nature of classification and the problems inherent in grouping the chaotic differences of the landscape. These include fuzzy set classification and "land characterization database" approaches. Fuzzy set classification of remotely sensed data incorporates uncertainty by valuing and displaying pixels based on the probability of membership in a given spectral group using a Beyesian or Dempster-Shafer procedure. In this way, contested membership in metaphoric land classes is reflected as a probability surface. This innovation attempts to incorporate uncertainty and ambiguity into planning and decision making for landscapes in which a unanimous interpretation does not exist. Even so, the technique can only reflect our uncertainty about the membership of any given pixel and location in any predefined set of land cover classes.

Land characteristics databases, as used in a prototype project at EROS data centre, are premised on the idea that land cover may be broken down into discrete, and seasonally distinct, elemental characteristics, including specific species mixes, ground surface cover, and canopy, which can then be combined to suit the needs of the analyst. These characteristics can, furthermore, be later aggregated to link local and regional studies to more global questions and to serve as a "rosetta stone" to translate between varying classification schemes. Nonetheless, while offering a departure from a single, arbitrary, classification technique, the land characteristics method is not free from the problems posed by more traditional techniques. The discrete elements that form the basis of the typology, be they the soil, hydrology, or canopy cover, are themselves identified before the fact. Classification is ultimately deductive then, rather than inductive; categories known in the mind's eye are found in the data, not vice versa. Even the most empirical clustering technique requires the naming and consolidation of clusters based on judgment or definition.

The data, however carefully differentiated, therefore, do not "speak for themselves," and more sensitive algorithms provide only an increasingly advanced and careful reproduction of existing categories, not the "discovery" of new ones.

In all cases then, the landscape vocabulary of the analyst is a necessary prerequisite to classification, and the analyst relies on an imaginary landscape against which to check resulting differentiation. This landscape is assembled from cover concepts that the analyst must already perceive based on features such as canopy, vegetation mixes, or apparent emptiness. Recognition of this fact is not new in debates over land cover classification. In consolidating and institutionalizing the USGS land use/land cover classification system, Anderson et al. emphasized that "there is no one ideal classification of land use and land cover... the process itself tends to be subjective even when an objective numerical approach is used." Anderson et al. also stressed that in aggregating and collapsing variable covers into a limited number of cover types, the interpreter must use subjective "patterns, tones, textures, and shapes" known before the fact.

Classification Metaphors are Cultural Artifacts

Unacknowledged, however, is the fact that such landscape "flags" and perceptions are culturally bound; as a classic example, an Inuit imagines and perceives the difference between upsik and api snow cover, where other observers might not. The larger conceptual language or "paradigm" of the observer drives the particular simplifications selected in analysis. The culture of forestry and remote sensing, into which a global population of analysts with similar training and background fall, ultimately produce and reproduce a common set of metaphors, and a common set of categories. Such categories reflect unproblematic classes that appear "universal" to the practitioners who share a common culture. These vary surprisingly little from case to case and reflect inherited notions of meaningful land classes. To illustrate, consider the range of categories conventionally deployed in remote sensing and land cover change analysis.

The most common cover categories, drawn from a survey of papers published in prominent international remote sensing journals between 1995 and 1999 include: Barren Land, Water, Grasses, Mixed Forest, Open Shrublands, Agricultural, Evergreen Broadleaf Forest, Forest Land, Wooded Grasslands/shrubs, and Urban or Built Up. These categories appear in a majority of the papers, and while the studies represent a range of geographical and ecological contexts, only a handful of categories dominate the analysis.

As Reed et al. (1994) note, "scientists typically select a land cover framework based on availability rather than suitability to a problem" or context. Similarly, the Anderson et al. (1976) classification itself, with its multi-tiered category system, allows a fine grade of distinctions, but its lower-level detailed categories are less often used in analysis. This occurs despite new techniques, such as image segmentation and filtering, which enable finer and more specific category sets.

This kind of uniformity in classification is an inherent and normal process of both daily experience and scientific analysis in a world of chaotic differences. There is nothing unscientific about the assignment of meaning that creates such categories of likeness. As in other arenas of "normal" science, consensus on such categories is necessary to order explanation and is part and prerequisite to analysis. At the same time, however, these associations, which define "which element pertains to which set," are fundamentally cultural and follow differing logic in differing contexts, both within and outside the remote sensing laboratory. Accepting even the most common and self-evident categories like "forest," "grasses," and "barren land" for the study of the distribution and change of landscapes without querying their socio-cultural sources or alternatives is problematic. The danger is not only that our analytical vocabulary is unnecessarily impoverished but, it is, moreover, that these culturally determined cover categories will inevitably reveal certain features, important to some observers, while disguising features that may well be important or relevant for other observers.

Categories as State Control

Competition between category sets is typically resolved through the imposition of a formal set of land cover classes over those of an already existing, in situ, cultural system. In the process, the imposed categories serve as a lever for control and prescription, and often act to reproduce the patterns of degradation and landscape change that rational and scientific planning seeks to halt. This problem occurs, to some degree, in the U.S. and many other developed world contexts; the culturally bound categories of the analyst or expert often contradict and confront those of the residents or communities for which images and maps are being produced. Such conflicts are even more apparent in the colonial and post-colonial context, where foreign categorical systems have been deployed uniformly over existing ones.

Land Cover Control in Colonialism

Colonial mapping entailed implicit (and sometimes explicit) unification of diverse spaces into uniform territory and thus required coherent and consistent land cover categories. The resulting maps were products of scientific state land management which, although itself a product of European science, was initially implemented exclusively in the colonial periphery. In these areas, from as early as the 18th century, conservation forestry began to exert a tremendous influence on colonial states, with a power unknown in Europe. In British colonies, such authority was especially marked. Drawing on the work of Alexander von Humboldt and other geographers, scientific connections between deforestation, drought, and economic crisis justified complex and stringent systems for land cataloguing and management. The goal of scientific forestry was to preserve the resource for public good through stringent techniques of naming and recording.

More than simply applying names, however, land classification under colonialism resulted in a political mapping of control, which linked the formation of "what is"—land capacities and characteristics—to what "should be"—land potentials and optimizations. This "survey modality" fixed and bounded natural objects in terms of how they ought to be used. Through the practice of surveying and mapping, colonial power accumulated bodies of information, created standard land cover types, and established these as norms by which the landscape must be judged and managed. Land covers imposed control by eliminating competing understandings.

6

Characteristics of Electromagnetic Radiation (EMR)

Electromagnetic Radiation

Electromagnetic waves are energy transported through space in the form of periodic disturbances of electric and magnetic fields.

All electromagnetic waves travel through space at the same speed, $c = 2.99792458 \times 10^8$ m/s, commonly known as the speed of light. An electromagnetic wave is characterized by a frequency and a wavelength. These two quantities are related to the speed of light by the equation,

speed of light = frequency × wavelength

The frequency (and hence, the wavelength) of an electromagnetic wave depends on its source. There is a wide range of frequency encountered in our physical world, ranging from the low frequency of the electric waves generated by the power transmission lines to the very high frequency of the gamma rays originating from the atomic nuclei.

This wide frequency range of electromagnetic waves constitute the Electromagnetic Spectrum.

THE ELECTROMAGNETIC SPECTRUM

The electromagnetic spectrum can be divided into several wavelength (frequency) regions, among which only a narrow band from about 400 to 700 nm is visible to the human eyes. Note that there is no sharp boundary between these regions.

Wavelength units: 1 mm = 1000 μm; 1 μm = 1000 nm.

- Radio Waves: 10 cm to 10 km wavelength.
- Microwaves: 1 mm to 1 m wavelength. The microwaves are further divided into different frequency (wavelength) bands: (1 GHz = 10^9 Hz)
  - o P band: 0.3-1 GHz (30-100 cm)
  - o L band: 1-2 GHz (15-30 cm)
  - o S band: 2-4 GHz (7.5-15 cm)
  - o C band: 4-8 GHz (3.8-7.5 cm)
  - o X band: 8-12.5 GHz (2.4-3.8 cm)
  - o Ku band: 12.5-18 GHz (1.7-2.4 cm)
  - o K band: 18-26.5 GHz (1.1-1.7 cm)
  - o Ka band: 26.5-40 GHz (0.75-1.1 cm)
- Infrared: 0.7 to 300 μm wavelength. This region is further divided into the following bands:
  - o Near Infrared (NIR): 0.7 to 1.5 μm.
  - o Short Wavelength Infrared (SWIR): 1.5 to 3 μm.
  - o Mid Wavelength Infrared (MWIR): 3 to 8 μm.
  - o Long Wanelength Infrared (LWIR): 8 to 15 μm.
  - o Far Infrared (FIR): longer than 15 μm.

The NIR and SWIR are also known as the Reflected Infrared, referring to the main infrared component of the solar radiation reflected from the earth's surface. The MWIR and LWIR are the Thermal Infrared.

- Visible Light: This narrow band of electromagnetic radiation extends from about 400 nm (violet) to about 700 nm (red). The various colour components of the visible spectrum fall roughly within the following wavelength regions:
  - o Red: 610-700 nm

  - o Orange: 590-610 nm
  - o Yellow: 570-590 nm
  - o Green: 500-570 nm
  - o Blue: 450-500 nm
  - o Indigo: 430-450 nm
  - o Violet: 400-430 nm
- Ultraviolet: 3 to 400 nm
- X-Rays and Gamma Rays.

Photons

According to quantum physics, the energy of an electromagnetic wave is quantized, i.e. it can only exist in discrete amount. The basic unit of energy for an electromagnetic wave is called a photon. The energy E of a photon is proportional to the wave frequency f,

$E = hf$

where the constant of proportionality h is the Planck's Constant,

$h = 6.626 \times 10^{-34}$ J s.

Visible Region

The visible light portion of the electromagnetic spectrum ranges from 0.4 micrometers ("μm") (shorter wavelength, higher frequency) to 0.7 μm (longer wavelength, lower frequency). This is the frequency range of light that the human eye is sensitive to. Every object reflects, absorbs and transmits electromagnetic energy in the visible portion of the electromagnetic spectrum and also other non-visible frequencies.

Electromagnetic energy which completely passes through an object is referred to as transmittance. Our eyes receive the visible light reflected from an object. The three primary colours reflected from an object known as *additive primaries* are the blue, green and red wavelengths.

Primary colours cannot be formed by the combination of any other primary colours. Intermediate colours are formed when a combination of primary colours are reflected from an object. Magenta is a combination of reflected red and blue, cyan a combination of reflected blue and green, and yellow a combination of reflected red and green.

Colour film produces colours by using layers of dyes which filter out various colours. The three colours which absorb the primary colours, known as *subtractive primaries*, are magenta, cyan and yellow. Magenta absorbs green and reflects red an blue, cyan absorbs red and reflects blue and green, and yellow absorbs blue and reflects red and green. The absorption of all colours produces black. If no colour is absorbed then the film produces white.

Infrared Region

The non-visible infrared spectral region lies between the visible light and the microwave portion of the electromagnetic spectrum. The infrared region covers a wavelength range from.7 μm to 14 μm. This broad range of infrared wavelengths is further subdivided into two smaller infrared regions. Each of these regions exhibits very different characteristics. The infrared region closest to visible light contains two smaller bands labeled near infrared and short-wave infrared with wavelengths ranging from.7 μm to 1.1 μm, and from 1.1 μm to 3.0 μm respectively. These infrared regions exhibit many of the same optical characteristics as visible light. The sun is the primary source of infrared radiation, which is reflected from an object. Cameras used to capture images in the visible light spectrum can capture images in the near infrared region by using special infrared film.

The other infrared region with longer wavelengths ranging from 3.0 μm to 14.0 μm is composed of two smaller bands labeled mid-wave infrared and long-wave infrared with wavelengths ranging from 3.0 μm to 5.0 μm, and from 5.0 μm to 14.0 μm respectively. Objects generate and emit thermal infrared radiation thus these objects can be detected at night because they are not dependent on reflected infrared radiation from the sun. Remote sensors operating in this infrared wavelength range measure an object's temperature.

BLACK BODY RADIATION

Black-body radiation is the thermal electromagnetic radiation within or surrounding a body in thermodynamic equilibrium with its environment, or emitted by a black body (an idealized opaque, non-reflective body). It has a specific spectrum and intensity that depends only on the body's temperature, which is assumed for the sake of calculations and theory to be uniform and constant.

The thermal radiation spontaneously emitted by many ordinary objects can be approximated as black-body radiation. A perfectly insulated enclosure that is in thermal equilibrium internally contains black-body radiation and will emit it through a hole made in its wall, provided the hole is small enough to have negligible effect upon the equilibrium.

A black body at room temperature appears black, as most of the energy it radiates is infra-red and cannot be perceived by the human eye. Because the human eye cannot perceive light waves at lower frequencies, a black body, viewed in the dark at the lowest just faintly visible temperature, subjectively appears grey, even though its objective physical spectrum peak is in the infrared range. When it becomes a little hotter, it appears dull red. As its temperature increases further it becomes yellow, white, and ultimately blue-white.

Although planets and stars are neither in thermal equilibrium with their surroundings nor perfect black bodies, black-body radiation is used as a first approximation for the energy they emit. Black holes are near-perfect black bodies, in the sense that they absorb all the radiation that falls on them. It has been proposed that they emit black-body radiation (called Hawking radiation), with a temperature that depends on the mass of the black hole.

The term *black body* was introduced by Gustav Kirchhoff in 1860. Black-body radiation is also called thermal radiation, *cavity radiation*, *complete radiation* or *temperature radiation*.

But there was one problem that was hard to get a grip on, an apparently blatant violation of the equipartition of energy. Consider an oven with a small hole in the door, through which the radiation inside is observed. This oven can be heated until it's white hot.

The radiation inside is infrared at low temperatures, becoming visible light as the temperature increases. So, the oven's full of electromagnetic waves, satisfying Maxwell's wave equation, with boundary conditions at the walls of the oven, the electric field has to be essentially zero there, because the walls conduct currents.

Of course, the radiation originates in oscillating charges in the walls, using the same analysis of Maxwell's equations that gives the radiation form an antenna. Anyway, there is a set of standing wave modes of electromagnetic vibrations inside the oven, just a three-dimensional version of the series of allowed standing wave modes of vibration of a string fixed at both ends.

So, we should be able to find the energy density of these waves using the same ideas that worked pretty well for the specific heats of solids and gases, that is to say, assume there's kT of energy in each mode of vibration. (This is $\frac{1}{2}kT$ of kinetic energy, $\frac{1}{2}kT$ of potential energy for each independent direction of vibration.)

But this leads to disaster. The problem is that there are infinitely many modes of vibration of the electromagnetic field in an oven. There is no upper limit to the number of wiggles the wave can have between the walls. So, if we take kT in each mode, we deduce that the oven contains an infinite amount of energy, and radiates an infinite amount through our small hole.

This analysis gives no clue as to why the colour we see changes with temperature. Evidently, equipartition of energy isn't working in this case. There's only a finite amount of energy in the oven and at low temperatures there's no energy at all in the modes corresponding to visible light, although that changes as things get hotter.

In the 1890's, German experimentalists measured the energy density as a function of wavelength to great precision, it's called the black body radiation spectrum. A theorist, Planck, found a mathematical formula that fitted this curve exactly,

$$R_T(f)\,df = \frac{8\pi V h f^3 df / c^3}{e^{hf/kT} - 1}$$

He did not at first have any theoretical justification for this formula, but it was a very accurate fit to some very precise experiments for a suitable value of the constant *h*, which we discuss in a moment.

Factoring out the number of modes of oscillation in the frequency range *df*, Planck's formula gives the average energy per mode to be,

$$\frac{hf}{e^{hf/kT} - 1}$$

For low frequencies, $hf \ll kT$, this correctly gives kT per mode.

But, for higher frequencies it's clear that the oscillators are not getting their "fair share" kT of energy. Somehow, the oscillating charges in the walls are not radiating so much energy at the high frequencies.

The only way Planck could derive the formula theoretically was by making a weird assumption: he assumed that the oscillating charges in the walls could *not* just radiate energy continuously, as Maxwell's equations would predict (and as was known to be true for ordinary antennas) but were only allowed to radiate energy in chunks he called *quanta*.

Furthermore, the amount of energy in one quantum depended on the frequency of the oscillation, in fact linearly: for frequency *f*, the quantum has energy *hf*, where *h* is the constant introduced into the formula above, now known as Planck's constant. It follows that the oscillators themselves could only be oscillating with energies that form a ladder with steps *hf* apart, above some lowest energy which would be their energy at absolute zero temperature.

The formula follows if we assume the oscillating field component in the oven having frequency *f* can only have a whole number of quanta of energy, that is to say, its energy must be one of: 0, *hf*, 2*hf*, 3*hf*. The relative probability of it having energy *E* is $e^{-E/kT}$, then its relative probabilities of having energy 0, *hf*, 2*hf*, are in the ratio 1: $e^{-hf/kT}$: $e^{-2hf/kT}$, etc.

The actual probabilities are given by dividing these relative probabilities by the sum of all of them.

They clearly are the terms of a geometric series, so their sum is just $1/(1 - e^{-hf/kT})$.

So, to find the average energy in the oscillator, we take the possible energies 0, *hf*, 2*hf*, 3*hf*, and weight each of them with their probability of occurring, that is, we must find

$$0.1 + hf.e^{-hf/kT} + 2hf.e^{-2hf/kT} + ...,$$

and divide the sum by $1/(1 - e^{-hf/kT})$.

So, Planck's quantum assumption explains the observed black body radiation curve. It also gives a qualitative explanation of the change in colour of the radiated light as the temperature is increased.

The oscillators in the walls derive their energy from the heat vibrations of neighbouring molecules: typically, such a vibration has energy of order *kT*, with probabilities of more energy going down as $e^{-E/kT}$.

This means that if the potentially radiating oscillator can only absorb energies in quanta *hf*, if $kT \ll hf$, it will be very unlikely to absorb *any* energy, and therefore very unlikely to radiate. In the three-dimensional oven, the number of standing wave oscillations in a small frequency range Δf increases with *f* as f^2, so we find that the maximum radiation intensity occurs at a frequency *f* such that *hf* is of order *kT*.

Therefore, as the temperature increases, the frequency at which the most intense radiation occurs increases, and hence the colour moves from red to blue.

Rayleigh-Jeans-Law

The excitation amplitude of each mode obeys the equation of motion of a harmonic oscillator. Therefore, classically one expects that each of mode is in thermal equilibrium excute with a thermal energy kT accordingtothe equipartition theorem, where k is Boltzmann's constant with $k = 1.38062 \pm 6 \cdot 10^{-23}$J/K

If that is the case the spectral energy density is given by the Rayleigh-Jeans-Law

$$w(f) = \frac{1}{V}\frac{dN}{df}kT = \frac{8\pi}{c^3}f^2kT.$$

this law describes very well the black body radiation for frequencies $hf << kT$ but there is an arbitrary large deviation for high frequencies.

This formula can not be correct, because it predicts infinite energy density for the high frequency modes resulting in an "ultraviolet catastrophy", i.e. the electromagnetic field containsaninfinite amount of energy at thermal equilibrium.

Wien's Law

The high frequency or short wavelength region of the black body radiation was first empirically described by Wien's Law

$$w(f) = \frac{8\pi hf^3}{c^3}e^{-hf/kt}..$$

Wien's law is surprisingly close to Planck's law, however it slightly fails to correctly predicts the asympthotic behaviour at low frequencies or long wavelengths.

Planck's Law

In the winter of 1900, Max Planck found the correct law for the black body radiation by assuming that each oscillator can only exchange energy in discrete portions or quanta. We rederive it by assuming that each mode can only have the discrete energie values.

$E_s = s \cdot hf$, for $s = 0, 1, 2...$

Thus s is the number of energy quanta stored in the oscillator. If the oscillator is a mode of the electromagnetic field we call s the number of photons. For the probability ps, that the oscillator has the energy Es we assume a Boltzmann distribution

$$P_s = \frac{1}{Z}\exp\left(-\frac{E_s}{kT}\right) = \frac{1}{Z}\exp\left(-\frac{hf}{kT}s\right)$$

where Z is a normalization factor such that the total propability of the oscillator to have any of the allowed energy values is

$$\sum_{s=0}^{\infty} P_s = 1\,.$$

Note, due to the fact th∠e oscillator energy is proportional to the number of photons, the statistics are exponential statistics. We obtain for the normalization factor

$$Z = \sum_{s=0}^{\infty} \exp\left(-\frac{hf}{kT}s\right) = \frac{1}{1-\exp\left(\frac{hf}{kT}\right)}$$

which is also called the partition function. The photon statistics are then given by

$$p_s = \exp\left(-\frac{hf}{kT}\right) = \left[1-\exp\left(\frac{hf}{kT}\right)\right]^{-1}$$

or with $\beta = \frac{hf}{kT}$

$$p_s = \frac{1}{Z(\beta)}e^{-\beta s}, with\, Z(\beta) = \sum_{s=0}^{\infty} e^{-\beta s} = \frac{1}{1-e^{-\beta}}$$

Given the statistics of the photon number, we can compute moments of the probability distribution, such as the average number of photons in the mode

$$\left\langle S^1 \right\rangle = \sum_{s=0}^{\infty} S^1 p_s$$

This first moment of the photon statistics can be computed from the partition function, using the "trick"

$$\left\langle S^1 \right\rangle = \frac{1}{Z(\beta)} \frac{\partial^1}{\partial(-\beta)^1} Z(\beta) = Z(\beta)^{e-\beta}$$

which is

$$\langle S \rangle = \frac{1}{\exp\frac{hf}{kT} - 1} .$$

With the average photon number $\langle s \rangle$ we obtain for the average energy stored in the mode $\langle E_s \rangle = \langle s \rangle hf$, and the energy density in the frequency interval $[f, f + df]$ is then given by

$$w(f) = \left\langle E_s \right\rangle \frac{dN}{df}$$

With the density of modes from Eq. we find Planck's law for the black body radiation

$$w(f) = \frac{8\pi f^2}{c^3} \frac{hf}{\exp\frac{hf}{kT} - 1}$$

which was used to make the plots. In the limits of low and high frequencies i.e. $hf << kT$ and $hf >> kT$ respectively Planck's law asympthotically approaches the Rayleigh-Jeans law and Wien's law.

RADIATION LAW

There are three phenomena through which energy can be transmitted: electromagnetic radiation, conduction, and convection. Unlike conduction and convection, electro-magnetic waves need no material medium for transmission.

Thus, light and radio waves can travel through interplanetary and interstellar space from the sun and stars to the earth. Regardless of the frequency, wavelength, or method of propagation, electromagnetic waves travel at a speed of 3×10^{10} cm (186,272 mi) per second in a vacuum. All the components of the electromagnetic spectrum, regardless of frequency, also have in common the typical properties of wave motion, including diffraction and interference. The wavelengths range from millionths of a centimeter to many kilometers.

The wavelength and frequency of electromagnetic waves are important in determining heating effect, visibility, penetration, and other characteristics of the electromagnetic radiation.

Theory of Electromagnetic Radiation

British physicist James Clerk Maxwell laid out the theory of electromagnetic waves in a series of papers published in the 1860s. He analysed mathematically the theory of electromagnetic fields and predicted that visible light was an electromagnetic phenomenon. Physicists had known since the early 19th century that light is propagated as a transverse wave (a wave in which the vibrations move in a direction perpendicular to the direction of the advancing wave front).

They assumed, however, that the wave required some material medium for its transmission, so they postulated an extremely diffuse substance, called ether, as the unobservable medium. Maxwell's theory made such an assumption unnecessary, but the ether concept was not abandoned immediately, because it fit in with the Newtonian concept of an absolute space-time frame for the universe.

A famous experiment conducted by the American physicist Albert Abraham Michelson and the American chemist Edward Williams Morley in the late 19th century served to dispel the ether concept and was important in the development of the theory of relativity. This work led to the realization that the speed of electromagnetic radiation in a vacuum is an invariant.

Ultraviolet Radiation

Ultraviolet Radiation, electromagnetic radiation that has wavelengths in the range between 4000 angstrom units (Å), the wavelength of violet light, and 150 Å, the length of X-rays. Natural ultraviolet radiation is produced principally by the sun. Ultraviolet radiation is produced artificially by electric-arc lamps. Ultraviolet radiation is often divided into three categories based on wavelength, *UV-A*, *UV-B*, and *UV-C*.

In general shorter wavelengths of ultraviolet radiation are more dangerous to living organisms. *UV-A* has a wavelength from 4000 Å to about 3150 Å. *UV-B* occurs at wavelengths from about 3150 Å to about 2800 Å and causes sunburn; prolonged exposure to *UV-B* over many years can cause skin cancer. *UV-C* has wavelengths of about 2800 Å to 150 Å and is used to sterilize surfaces because it kills bacteria and viruses.

The earth's atmosphere protects living organisms from the sun's ultraviolet radiation. If all the ultraviolet radiation produced by the sun were allowed to reach the surface of the earth, most life on earth would probably be destroyed. Fortunately, the ozone layer of the atmosphere absorbs almost all of the short-wavelength ultraviolet radiation, and much of the long-wavelength ultraviolet radiation. However, ultraviolet radiation is not entirely harmful; a large portion of the vitamin *D* that humans and animals need for good health is produced when the human's or animal's skin is irradiated by ultraviolet rays.

When exposed to ultraviolet light, many substances behave differently than when exposed to visible light. For example, when exposed to ultraviolet radiation, certain minerals, dyes, vitamins, natural oils, and other products become *fluorescent*—that is, they appear to glow. Molecules in the substances absorb the invisible ultraviolet light, become energetic, then shed their excess energy by emitting visible light.

As another example, ordinary window glass, transparent to visible light, is opaque to a large portion of ultraviolet rays, particularly ultraviolet rays with short wavelengths. Special-formula glass is transparent to the longer ultraviolet wavelengths, and quartz is transparent to the entire naturally occurring range.

INTERACTION OF EMR WITH OF ATMOSPHERE

When electromagnetic radiation travels through the atmosphere, it may be absorbed or scattered by the constituent particles of the atmosphere. Molecular absorption converts the radiation energy into excitation energy of the molecules. Scattering redistributes the energy of the incident beam to all directions. The overall effect is the removal of energy from the incident radiation. The various effects of absorption and scattering are outlined in the following sections.

Atmospheric Transmission Windows

Each type of molecule has its own set of absorption bands in various parts of the electromagnetic spectrum. As a result, only the wavelength regions outside the main absorption bands of the atmospheric gases can be used for remote sensing. These regions are known as the Atmospheric Transmission Windows. The wavelength bands used in remote sensing systems are usually designed to fall within these windows to minimize the atmospheric absorption effects. These windows are found in the visible, near-infrared, certain bands in thermal infrared and the microwave regions.

Effects of Atmospheric Absorption

Atmospheric absorption affects mainly the visible and infrared bands. Optical remote sensing depends on solar radiation as the source of illumination. Absorption reduces the solar radiance within the absorption bands of the atmospheric gases. The reflected radiance is also attenuated after passing through the atmosphere. This attenuation is wavelength dependent. Hence, atmospheric absorption will alter the apparent spectral signature of the target being observed.

Effects of Atmospheric Scattering

Atmospheric scatterring is important only in the visible and near infrared regions. Scattering of radiation by the constituent gases and aerosols in the atmosphere causes degradation of the remotely sensed images. Most noticeably, the solar radiation scattered by the atmosphere towards the sensor without first reaching the ground produces a hazy appearance of the image. This effect is particularly severe in the blue end of the visible spectrum due to the stronger Rayleigh Scattering for shorter wavelength radiation.

Furthermore, the light from a target outside the field of view of the sensor may be scattered into the field of view of the sensor. This effect is known as the adjacency effect. Near to the boundary between two regions of different brightness, the adjacency effect results in an increase in the apparent brightness of the darker region while the apparent brightness of the brighter region is reduced. Scattering also produces blurring of the targets in remotely sensed images due to spreading of the reflected radiation by scattering, resulting in a reduced resolution image.

Airborne Remote Sensing

In airborne remote sensing, downward or sideward looking sensors are mounted on an aircraft to obtain images of the earth's surface. An advantage of airborne remote sensing, compared to satellite remote sensing, is the capability of offering very high spatial resolution images (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, whereas earth observation satellites offer the possibility of continuous monitoring of the earth.

Analog aerial photography, videography, and digital photography are commonly used in airborne remote sensing. Synthetic Aperture Radar imaging is also carried out on airborne platforms.

Analog photography is capable of providing high spatial resolution. The interpretation of analog aerial photographs is usually done visually by experienced analysts. The photographs may be digitized using a scanning device for computer-assisted analysis.

Digital photography permits real-time transmission of the remotely sensed data to a ground station for immediate analysis. The digital images can be analysed and interpreted with the aid of a computer.

In spaceborne remote sensing, sensors are mounted on-board a spacecraft (space shuttle or satellite) orbiting the earth. At present, there are several remote sensing satellites providing imagery for research and operational applications. Spaceborne 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;
- Semiautomated computerised processing and analysis;
- Relatively lower cost per unit area of coverage.

Satellite imagery has a generally lower resolution compared to aerial photography. However, very high resolution imagery (up to 1-m resolution) is now commercially available to civilian users with the successful launch of the IKONOS-2 satellite in September 24, 1999.

Satellite Orbits

A satellite follows a generally elliptical orbit around the earth. The time taken to complete one revolution of the orbit is called the orbital period. The satellite traces out a path on the earth surface, called its ground track, as it moves across the sky. As the earth below is rotating, the satellite traces out a different path on the ground in each subsequent cycle. Remote sensing satellites are often launched into special orbits such that the satellite repeats its path after a fixed time interval. This time interval is called the repeat cycle of the satellite.

EARTH'S SURFACE REFLECTION, ABSORPTION, TRANSMISSION, SCATTERING AND REFRACTION

The reflecting power of a surface is known as 'albedo'. Bright snow and ice have a high albedo, meaning they reflect solar radiation back into space, while green areas like forests and fields have a much lower albedo.

The lower the albedo, the more energy from the Sun is absorbed.

Changes in Earth's surfaces can therefore affect how much of the Sun's energy is absorbed – such as a decrease in snow cover or an increase in the area used for agriculture. If the amount of energy absorbed changes, this has an effect on Earth's energy budget and ultimately affects our weather and climate.

To help scientists build better simulations of weather and climate, ESA's GlobAlbedo project is using satellite data to map changes in Earth's reflectivity.

Led by University College London, the team used readings from the Envisat and Spot-Vegetation satellites to produce global surface albedo maps from 1998 to 2011. The maps, available for free online, provide the most accurate measure of Earth's reflectivity to date.

"GlobAlbedo is the first gap-free, 1 km-resolution map of Earth's land surface with an uncertainty estimate for every pixel. This could only have been produced from satellite data," said Professor Jan-Peter Muller of University College London, leader of the GlobAlbedo project.

By combining data from different satellite sensors, scientists have maximised the coverage and created a time series that can be extended to include historical as well as future satellite measurements.

The maps have proven useful to a variety of users, including the UK Met Office. Scientists there have been using them to update the land surface albedo information in the Met Office's operational Global Atmosphere weather model, resulting in more accurate weather predictions and climate forecasts.

"Tests show that they help to give more accurate temperature forecasts over the United States and Asia, especially in summer," said Dr Malcolm Brooks from the Met Office. "We expect to be producing operational forecasts using these data in the spring of 2014."

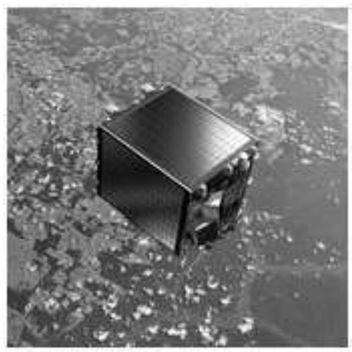


Fig. Proba-V satellite

Other case-studies have been looking into different uses of land albedo. These include investigating the effect of anthropogenic changes in land cover on Earth's energy balance, studying how agricultural practices influence heat waves and verifying climate models.

The mapping and monitoring of Earth's albedo will continue with ESA's recently launched Proba-V satellite and the future Sentinel-3 mission being developed under Europe's Global Monitoring for Environment and Security programme, Copernicus.

Absorption (electromagnetic radiation)

In physics, absorption of electromagnetic radiation is how matter (typically electrons bound in atoms) takes up a photon's energy — and so transforms electromagnetic energy into internal energy of the absorber (for example, thermal energy). A notable effect (attenuation) is to gradually reduce the intensity of light waves as they propagate through a medium. Although the absorption of waves does not usually depend on their intensity (linear absorption), in certain conditions (optics) the medium's transparency changes by a factor that varies as a function of wave intensity, and saturable absorption (or nonlinear absorption) occurs.

Quantifying absorption

Many approaches can potentially quantify radiation absorption, with key examples following.

- The absorption coefficient along with some closely related derived quantities
- The attenuation coefficient (NB used infrequently with meaning synonymous with "absorption coefficient")
- The molar attenuation coefficient (also called "molar absorptivity"), which is the absorption coefficient divided by molarity
- The mass attenuation coefficient (also called "mass extinction coefficient"), which is the absorption coefficient divided by density
- The absorption cross section and scattering cross-section, related closely to the absorption and attenuation coefficients, respectively
- "Extinction" in astronomy, which is equivalent to the attenuation coefficient
- Other measures of radiation absorption, including penetration depth and skin effect, propagation constant, attenuation constant, phase constant, and complex wavenumber, complex refractive index and extinction coefficient, complex dielectric constant, electrical resistivity and conductivity.
- Related measures, including absorbance (also called "optical density") and optical depth (also called "optical thickness")

All these quantities measure, at least to some extent, how well a medium absorbs radiation. Which among them practitioners use varies by field and technique, often due simply to convention.

Measuring absorption

The absorbance of an object quantifies how much of the incident light is absorbed by it (instead of being reflected or refracted). This may be related to other properties of the object through the Beer–Lambert law.

Precise measurements of the absorbance at many wavelengths allow the identification of a substance via absorption spectroscopy, where a sample is illuminated from one side, and the intensity of the light that exits from the sample in every direction is measured.

A few examples of absorption are ultraviolet–visible spectroscopy, infrared spectroscopy, and X-ray absorption spectroscopy.

Applications

Understanding and measuring the absorption of electromagnetic radiation has a variety of applications.

- In radio propagation, it is represented in non-line-of-sight propagation. For example, see computation of radio wave attenuation in the atmosphere used in satellite link design.
- In meteorology and climatology, global and local temperatures depend in part on the absorption of radiation by atmospheric gases (such as in the greenhouse effect) and land and ocean surfaces.
- In medicine, X-rays are absorbed to different extents by different tissues (bone in particular), which is the basis for X-ray imaging.
- In chemistry and materials science, different materials and molecules absorb radiation to different extents at different frequencies, which allows for material identification.
- In optics, sunglasses, colored filters, dyes, and other such materials are designed specifically with respect to which visible wavelengths they absorb, and in what proportions.
- In biology, photosynthetic organisms require that light of the appropriate wavelengths be absorbed within the active

area of chloroplasts, so that the light energy can be converted into chemical energy within sugars and other molecules.

- In physics, the D-region of Earth's ionosphere is known to significantly absorb radio signals that fall within the high-frequency electromagnetic spectrum.
- In nuclear physics, absorption of nuclear radiations can be used for measuring the fluid levels, densitometry or thickness measurements.

Reflection, Transmission, and Absorption

Reflection is the process by which electromagnetic radiation is returned either at the boundary between two media (surface reflection) or at the interior of a medium (volume reflection), whereas transmission is the passage of electromagnetic radiation through a medium. Both processes can be accompanied by diffusion (also called scattering), which is the process of deflecting a unidirectional beam into many directions. In this case, we speak about diffuse reflection and diffuse transmission . When no diffusion occurs, reflection or transmission of an unidirectional beam results in an unidirectional beam according to the laws of geometrical optics. In this case, we speak about regular reflection (or specular reflection) and regular transmission (or direct transmission). Reflection, transmission and scattering leave the frequency of the radiation unchanged. Exception: The Doppler effect causes a change in frequency when the reflecting material or surface is in motion.

Absorption is the transformation of radiant power to another type of energy, usually heat, by interaction with matter.

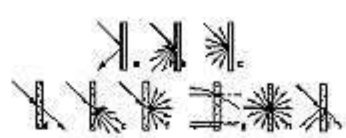


Fig: Direct, mixed and diffuse reflection d-f: direct, mixed and diffuse transmission

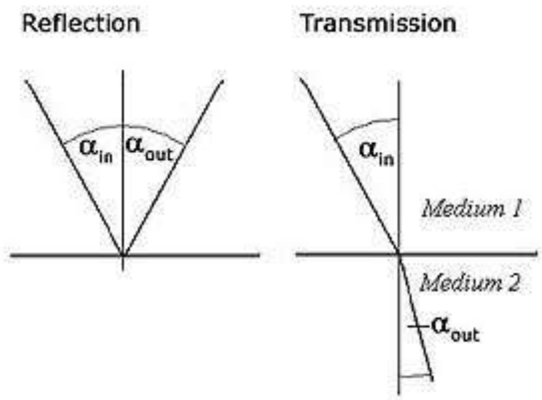
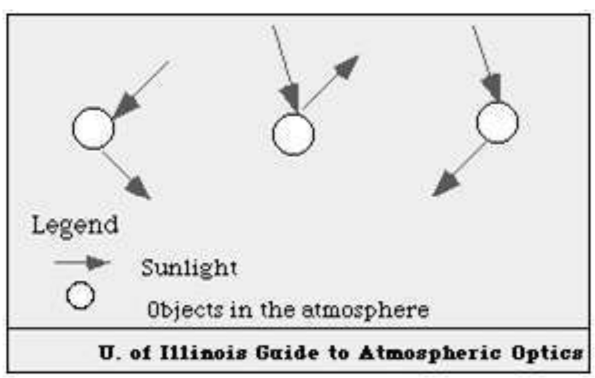


Fig. : When directly reflected or directly transmitted, an unidirectional beam follows the laws of geometrical optics

Tadiation scattering

Radiation scattering: The diversion of radiation (thermal, electromagnetic, or nuclear) from its original path as a result of interaction or collisions with atoms, molecules, or larger particles in the atmosphere or other media between the source of radiation (*e.g.*, a nuclear explosion) and a point some distance away. As a result of scattering, radiation (especially gamma rays and neutrons) will be received at such a point from many directions instead of only from the direction of the source.



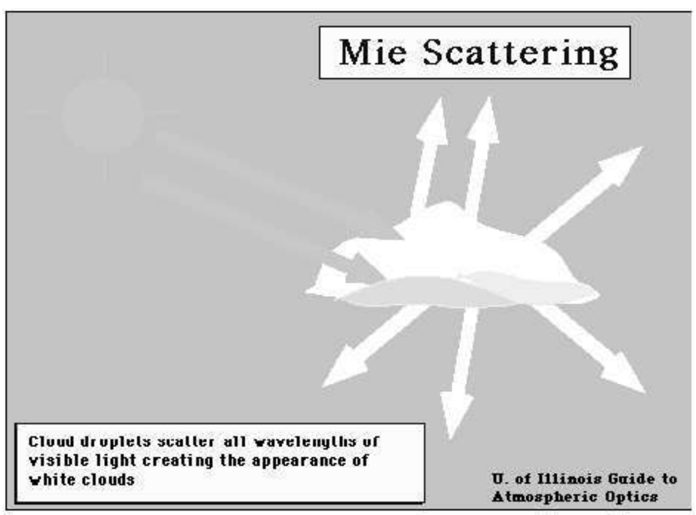
Scattering is the process by which "small particles suspended in a medium of a different index of refraction diffuse a portion of the incident radiation in all directions." With scattering, there is no energy transformation, but a change in the spatial distribution of the energy. Scattering, along with absorption, causes attenuation problems with radar and other measuring devices.

In the above graphic by the University of Illinois, sunlight comes into the atmosphere and can be scattered in any direction as it passes through a medium. This diffuses the light— spreading it out in all directions so it is not just a single, straight beam. If it was not for scattering, we would not be able to see shadowed objects such as walnuts that have fallen on the ground under the shading of a tree.

There are three different types of scattering: Rayleigh scattering, Mie scattering, and non-selective scattering.

Rayleigh scattering mainly consists of scattering from atmospheric gases. This occurs when the particles causing the scattering are smaller in size than the wavelengths of radiation in contact with them. This type of scattering is therefore wavelength dependent. As the wavelength decreases, the amount of scattering increases. Because of Rayleigh scattering, the sky appears blue, as in the picture below. This is because blue light is scattered around four times as much as red light, and UV light is scattered about 16 times as much as red light. Mie scattering is caused by pollen, dust, smoke, water droplets, and other particles in the lower portion of the atmosphere. It occurs when the particles causing the scattering are larger than the wavelengths of radiation in contact with them. Mie scattering is responsible for the white appearance of the clouds, as seen below.

The effects are also wavelength dependent. A graphic about mie scattering from the University of Illinois is shown below.



The last type of scattering is non-selective scattering. It occurs in the lower portion of the atmosphere when the particles are much larger than the incident radiation. This type of scattering is not wavelength dependent and is the primary cause of haze.

Refraction

In physics refraction is the change in direction of a wave passing from one medium to another or from a gradual change in the medium.Refraction of light is the most commonly observed phenomenon, but other waves such as sound waves and water waves also experience refraction. How much a wave is refracted is determined by the change in wave speed and the initial direction of wave propagation relative to the direction of change in speed.

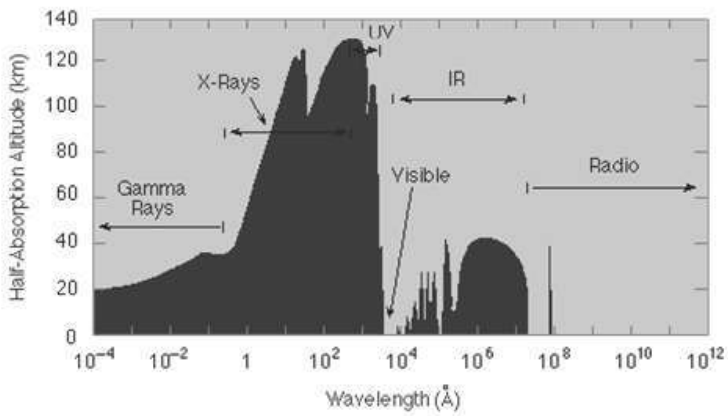
General explanation

Consider a wave going from one material to another where its speed is slower as in the figure. If it reaches the interface between the materials at an angle one side of the wave will reach the second material first, and therefore slow down earlier. With one side of the wave going slower the whole wave will pivot towards that side. This is why a wave will bend away from the surface or toward the normal when going into a slower material. In the opposite case of a wave reaching a material where the speed is higher, one side of the wave will speed up and the wave will pivot away from that side.

ATMOSPHERIC WINDOWS

One important practical consequence of the interaction of electromagnetic radiation with matter and of the detailed composition of our atmosphere is that only light in certain wavelength regions can penetrate the atmosphere well. These regions are called *atmospheric windows*.

The following figure shows the amount of absorption at different wavelengths in the atmosphere. It is presented in terms of the *half-absorption altitude*, which is defined to be the altitude in the atmosphere (measured from the Earth's surface) where 1/2 of the radiation of a given wavelength incident on the upper atmosphere has been absorbed. Windows correspond to those regions where the half-absorption altitude is very small.



The dominant windows in the atmosphere are seen to be in the visible and radio frequency regions, while X-Rays and UV are seen to be very strongly absorbed and Gamma Rays and IR are somewhat less strongly absorbed. We see clearly the argument for getting above the atmosphere with detectors on space-borne platforms in order to observe at wavelengths other than the visible and RF regions.

The Atmospheric Window as a Curtain Filled of Holes

Not all of the EM spectrum hits the Earth's surface. Atmospheric absorption prevents specific types of EM radiation to pass through the atmosphere.

The upper atmosphere blocks 100% of the gamma rays, x-rays and most ultra-violet light. But notice how visible light freely passes. Our eyes use this visible light to see features on Earth.

Think of the atmospheric window like a curtain with holes. And specific multispectral and hyperspectral bands can freely pass.

The curtain (atmospheric window) blocks specific wavelengths like gamma rays. But in reality, the curtain is the water vapor, ozone, carbon dioxide and molecules in the atmosphere. These blocked wavelengths are known as "absorption bands".

Holes in the curtain are the "atmospheric window" with specific bands of EM spectrum can freely pass. In the image below, you can see how visible and near infrared light almost pass freely.

To recap, an atmospheric window is the portion of the electromagnetic spectrum that can be transmitted through the atmosphere. And an absorption band is the portion of the electromagnetic spectrum that can be transmitted through the atmosphere.

How the Atmospheric Window Impacts Remote Sensing

The atmospheric window allows specific types of EM radiation to freely pass. Radio waves can pass through quite easily. But x-rays cannot and are blocked in the absorption band.

7

Fundamentals of Aerial Photography

INTRODUCTION

Aerial photography is the taking of photographs of the ground from an elevated position. The term usually refers to images in which the camera is not supported by a ground-based structure. Cameras may be hand held or mounted, and photographs may be taken by a photographer, triggered remotely or triggered automatically. Platforms for aerial photography include fixed-wing aircraft, helicopters, balloons, blimps and dirigibles, rockets, kites, poles, parachutes,vehicle mounted poles. Aerial photography should not be confused with Air-to-Air Photography, when aircraft serve both as a photo platform and subject.

History

Honoré Daumier, "Nadar élevant la Photographie à la hauteur de l'Art" (Nadar elevating Photography to Art), published in *Le Boulevard*, May 25, 1862.

Aerial photography was first practiced by the French photographer and balloonist Gaspard-Félix Tournachon, known as "Nadar", in 1858 over Paris, France.

The first use of a motion picture camera mounted to a heavier-than-air aircraft took place on April 24, 1909 over Rome in the 3:28 silent film short, *Wilbur Wright und seine Flugmaschine*. The first special semiautomatic aerial camera was designed in 1911 by Russian military engineer — Colonel Potte V. F. This aerial camera was used during World War I.

The use of aerial photography for military purposes was expanded during World War I by many other aviators such as Fred Zinn. One of the first notable battles was that of Neuve Chapelle.

Aerial mapping came into use on the battlefronts during World War I. In January 1918, General Allenby used five Australian pilots from No. 1 Squadron AFC to photograph a 624 square miles (1,620 km^2) area in Palestine as an aid to correcting and improving maps of the Turkish front. Lieutenants Leonard Taplin, Allan Runciman Brown, H. L. Fraser, Edward Patrick Kenny, and L. W. Rogers photographed a block of land stretching from the Turkish front lines 32 miles (51 km) deep into their rear areas. Beginning 5 January, they flew with a fighter escort to ward off enemy fighters. Using Royal Aircraft Factory BE.12 and Martinsyde airplanes, they not only overcame enemy air attacks, but also bucked 65 mile per hour winds, antiaircraft fire, and malfunctioning equipment to complete their task circa 19 January 1918.

One of the most successful pioneers of the commercial use of aerial photography was by Sherman Fairchild who started his own aircraft firm Fairchild Aircraft to develop and build specialized aircraft for high altitude aerial survey missions. One Fairchild aerial survey aircraft in 1935 carried unit that combined two synchronized cameras, and each camera having five six inch lenses with a ten inch lenses and took photos from 23,000 feet. Each photo cover two hundred and twenty five square miles. One of its first government contracts was an aerial survey of New Mexico to study soil erosion. A year later, Fairchild introduced a better high altitude camera with nine-lens in one unit that could take a photo of 600 square miles with each exposure from 30,000 feet.

With the advent of inexpensive digital cameras, many people now take candid photographs from commercial aircraft and increasingly from general aviation aircraft on private pleasure flights.

Uses of Imagery

Aerial photography is used in cartography (particularly in photogrammetric surveys, which are often the basis for topographic maps), land-use planning, archaeology, movie production, environmental studies, surveillance, commercial advertising, conveyancing, and artistic projects. In the United States, aerial photographs are used in many Phase I Environmental Site Assessments for property analysis. Aerial photos are often processed using GIS software.

Infrared Photography

In infrared photography, the film or image sensor used is sensitive to infrared light. The part of the spectrum used is referred to as near-infrared to distinguish it from far-infrared, which is the domain of thermal imaging. Wavelengths used for photography range from about 700 nm to about 900 nm. Usually an "infrared filter" is used; this lets infrared (IR) light pass through to the camera, but blocks all or most of the visible light spectrum (the filter thus looks black or deep red). When these filters are used together with infrared-sensitive film or sensors, very interesting "in-camera effects" can be obtained; false-colour or black-and-white images with a dreamlike or sometimes lurid appearance known as the "Wood Effect," an effect mainly caused by foliage (such as tree leaves and grass) strongly reflecting in the same way visible light is reflected from snow. There is a small contribution from chlorophyll fluorescence, but this is marginal and is not the real cause of the brightness seen in infrared photographs. The effect is named after the infrared photography pioneer Robert W. Wood, and not after the material wood, which does not strongly reflect infrared.

The other attributes of infrared photographs include very dark skies and penetration of atmospheric haze, caused by reduced Rayleigh scattering and Mie scattering, respectively, compared to visible light. The dark skies, in turn, result in less infrared light in shadows and dark reflections of those skies from water, and clouds will stand out strongly. These wavelengths also penetrate a few millimeters into skin and give a milky look to portraits, although eyes often look black.

History

Until the early 20th century, infrared photography was not possible because silver halide emulsions are not sensitive to longer wavelengths than that of blue light (and to a lesser extent, green light) without the addition of a dye to act as a colour sensitizer. The first infrared photographs (as distinct from spectrographs) to be published appeared in the February 1910 edition of The Century Magazine and in the October 1910 edition of the Royal Photographic Society Journal to illustrate papers by Robert W. Wood, who discovered the unusual effects that now bear his name.

The RPS is co-ordinating events to celebrate the centenary of this event in 2010. Wood's photographs were taken on experimental film that required very long exposures; thus, most of his work focused on landscapes. A further set of infrared landscapes taken by Wood in Italy in 1911 used plates provided for him by CEK Mees at Wratten & Wainwright. Mees also took a few infrared photographs in Portugal in 1910, which are now in the Kodak archives.

Infrared-sensitive photographic plates were developed in the United States during World War I for spectroscopic analysis, and infrared sensitizing dyes were investigated for improved haze penetration in aerial photography. After 1930, new emulsions from Kodak and other manufacturers became useful to infrared astronomy.

Infrared photography became popular with photography enthusiasts in the 1930s when suitable film was introduced commercially. The Times regularly published landscape and aerial photographs taken by their staff photographers using Ilford infrared film. By 1937 33 kinds of infrared film were available from five manufacturers including Agfa, Kodak and Ilford. Infrared movie film also available and was used to create day-for-night effects in motion pictures, a notable example being the pseudo-night aerial sequences in the James Cagney/Bette Davis movie The Bride Came COD.

False-colour infrared photography became widely practiced with the introduction of Kodak Ektachrome Infrared Aero Film and Ektachrome Infrared EIR. The first version of this, known as Kodacolor Aero-Reversal-Film, was developed by Clark and others at the Kodak for camouflage detection in the 1940s. The film became more widely available in 35mm form in the 1960s but Kodak Aerochrome III Infrared Film 1443 is their sole remaining infrared film.

Infrared photography became popular with a number of 1960s recording artists, because of the unusual results; Jimi Hendrix, Donovan, Frank Zappa and the Grateful Dead all issued albums with infrared cover photos. The unexpected colours and effects that infrared film can produce fit well with the psychedelic aesthetic that emerged in the late 1960s.

For some, infrared photography can easily look gimmicky, but many photographers such as Elio Ciol and Martin Reeves have made subtle use of black-and-white infrared-sensitive film. With the advent of digital infrared photography, as a part of full spectrum photography, the technique is gaining popularity and is being sold as fine art photographs in a variety of galleries worldwide.

Infrared light lies between the visible and microwave portions of the electromagnetic spectrum. Infrared light has a range of wavelengths, just like visible light has wavelengths that range from red light to violet. "Near infrared" light is closest in wavelength to visible light and "far infrared" is closer to the microwave region of the electromagnetic spectrum. The longer, far infrared wavelengths are about the size of a pin head and the shorter, near infrared ones are the size of cells, or are microscopic.

Focusing Infrared

Most manual focus 35 mm SLR and medium format SLR lenses have a red dot, line or diamond, often with a red "R" called the infrared index mark, that can be used to achieve proper infrared focus; many autofocus lenses no longer have this mark. When a single-lens reflex (SLR) camera is fitted with a filter that is opaque to visible light, the reflex system becomes useless for both framing and focusing, one must compose the picture without the filter and then attach the filter. This requires the use of a tripod to prevent the composition from changing. A sharp infrared photograph can be done with a tripod, a narrow aperture (like *f*/22) and a slow shutter speed without focus compensation, however wider apertures like *f*/2.0 can produce sharp photos only if the lens is meticulously refocused to the infrared index mark, and only if this index mark is the correct one for the filter and film in use. However, it should be noted that diffraction effects inside a camera are greater at infrared wavelengths so that stopping down the lens too far may actually reduce sharpness.

Most apochromatic ('APO') lenses do not have an Infrared index mark and do not need to be refocused for the infrared spectrum because they are already optically corrected into the near-infrared spectrum. Catadioptric lenses do not require this adjustment because mirrors do not suffer from chromatic aberration. Zoom lenses may scatter more light through their more complicated optical systems than prime lenses, that is, lenses of fixed focal length; for example, an infrared photo taken with a 50 mm prime lens may look more contrasty than the same image taken at 50 mm with a 28–80 zoom.

Some lens manufacturers such as Leica never put IR index marks on their lenses. The reason for this is because any index mark is only valid for one particular IR filter and film combination, and may lead to user error. Even when using lenses with index marks, focus testing is advisable as there may be a large difference between the index mark and the subject plane.

Film Cameras

Many conventional cameras can be used for infrared photography, where infrared is taken to mean light of a wavelength only slightly longer than that of visible light. Photography of rather longer wavelengths is normally termed thermography and requires special equipment.

With some patience and ingenuity, most film cameras can be used. However, some cameras of the 1990s that used 35mm film have infrared sprocket-hole sensors that can fog infrared film (their manuals may warn against the use of infrared film for this reason). Other film cameras are not completely opaque to infrared light.

Black-and-white Infrared Film

Black-and-white infrared negative films are sensitive to wavelengths in the 700 to 900 nm near infrared spectrum, and most also have a sensitivity to blue light wavelengths.

The notable halation effect or glow often seen in the highlights of infrared photographs is an artifact of Kodak High Speed Infrared (HIE) black-and-white negative film and not an artifact of infrared light. The glow or blooming is caused by the absence of an anti-halation layer on the back side of Kodak HIE film, this results in a scattering or blooming around the highlights that would usually be absorbed by the anti-halation layer in conventional films.

The majority of black-and-white infrared art, landscape, and wedding photography is done using orange (15 or 21), red (23, 25, or 29) or visually opaque (72) filters over the lens to block the blue visible light from the exposure.

The intent of filters in black-and-white infrared photography is to block blue wavelengths and allow infrared to pass through. Without filters, infrared negative films look much like conventional negative films because the blue sensitivity lowers the contrast and effectively counteracts the infrared look of the film. Some photographers use orange or red filters to allow slight amounts of blue wavelengths to reach the film, and thus lower the contrast. Very dark-red (29) filters block out almost all blue, and visually opaque (70, 89b, 87c, 72) filters block out all blue and also visible-red wavelengths, resulting in a more pure-infrared photo with a more pronounced contrast.

Certain infrared-sensitive films like Kodak HIE must only be loaded and unloaded in total darkness. Infrared black-and-white films require special development times but development is usually achieved with standard black-and-white film developers and chemicals (like D-76). Kodak HIE film has a polyester film base that is very stable but extremely easy to scratch, therefore special care must be used in the handling of Kodak HIE throughout the development and printing/scanning process to avoid damage to the film.

As of November 2, 2007, "KODAK is preannouncing the discontinuance" of HIE Infrared 35 mm film stating the reasons that, "Demand for these products has been declining significantly in recent years, and it is no longer practical to continue to manufacture given the low volume, the age of the product formulations and the complexity of the processes involved." At the time of this notice, HIE Infrared 135-36 was available at a street price of around $12.00 a roll at US mail order outlets.

Arguably the greatest obstacle to infrared film photography has been the increasing difficulty of obtaining infrared-sensitive film. However, despite the discontinuance of HIE, other newer infrared sensitive emulsions from EFKE, ROLLEI, and ILFORD are still available, but these formulations have differing sensitivity and specifications from the venerable KODAK HIE that has been around for at least two decades. Some of these infrared films are available in 120 and larger formats as well as 35 mm, which adds flexibility to their application. With the discontinuance of Kodak HIE, Efke's IR820 film has become the only IR film on the market with good sensitivity beyond 750 nm, the Rollei film does extend beyond 750 nm but IR sensitivity falls off very rapidly.

Colour Infrared Film

Colour infrared transparency films have three sensitized layers that, because of the way the dyes are coupled to these layers, reproduce infrared as red, red as green, and green as blue. All three layers are sensitive to blue so the film must be used with a yellow filter, since this will block blue light but allow the remaining colours to reach the film.

The health of foliage can be determined from the relative strengths of green and infrared light reflected; this shows in colour infrared as a shift from red (healthy) towards magenta (unhealthy). Early colour infrared films were developed in the older E-4 process, but Kodak later manufactured a colour transparency film that could be developed in standard E-6 chemistry, although more accurate results were obtained by developing using the AR-5 process. In general, colour infrared does not need to be loaded in total darkness (despite what it says on the can), or refocused to the infrared index mark on the lens.

In 2007 Kodak announced that production of the 35 mm version of their colour infrared film (Ektachrome Professional Infrared/ EIR) would cease as there was insufficient demand. It is assumed that the 70 mm Aerographic format will continue.

There is no currently available digital camera that will produce the same results as Kodak colour infrared film although the equivalent images can be produced by taking two exposures, one infrared and the other full-colour, and combining in post-production. The colour images produced by digital still cameras using infrared-pass filters are not equivalent to those produced on colour infrared film. The colours result from varying amounts of infrared passing through the colour filters on the photo sites, further amended by the Bayer filtering. While this makes such images unsuitable for the kind of applications for which the film was used, such as remote sensing of plant health, the resulting colour tonality has proved popular artistically.

Colour digital infrared, as part of full spectrum photography is gaining popularity. The ease of creating a softly coloured photo with infrared characteristics has found interest among hobbyists and professionals, with art work being sold in galleries since 2005.

Availability

Kodak colour infrared film for 35 mm has been discontinued. It is still available in 120 medium format from a supplier in Germany who cuts it down from fresh bulk stock. Otherwise, there is no known supplier of fresh stock. The 35 mm rolls that are still available on various sites, are for the most part, expired. (As of February 2010, the supplier in Germany announced that the film he has on hand will be the last- Kodak has discontinued the large roll stock he was using to cut down to 120 and sheet formats.)

Digital Cameras

Digital camera sensors are inherently sensitive to infrared light, which would interfere with the normal photography by confusing the autofocus calculations or softening the image (because infrared light is focused differently than visible light), or oversaturating the red channel. Also, some clothing is transparent in the infrared, leading to unintended (at least to the manufacturer) uses of video cameras. Thus, to improve image quality and protect privacy, many digital cameras employ infrared blockers. Depending on the subject matter, infrared photography may not be practical with these cameras because the exposure times become overly long, often in the range of 30 seconds, creating noise and motion blur in the final image. However, for some subject matter the long exposure does not matter or the motion blur effects actually add to the image. Some lenses will also show a 'hot spot' in the centre of the image as their coatings are optimised for visible light and not for IR.

An alternative method of DSLR infrared photography is to remove the infrared blocker in front of the charge-coupled device and replace it with a filter that removes visible light. This filter is behind the mirror, so the camera can be used normally- handheld, normal shutter speeds, normal composition through the viewfinder, and focus, all work like a normal camera. Metering works but is not always accurate because of the difference between visible and infrared reflection. When the IR blocker is removed, many lenses which did display a hotspot cease to do so, and become perfectly usable for infrared photography. Additionally, because the red, green and blue micro-filters remain and have transmissions not only in their respective colour but also in the infrared, enhanced infrared colour may be recorded.

Since the Bayer filters in most digital cameras absorb a significant fraction of the infrared light, these cameras are sometimes not very sensitive as infrared cameras and can sometimes produce false colours in the images. An alternative approach is to use a Foveon X3 sensor, which does not have absorptive filters on it; the Sigma SD10 DSLR has a removable IR blocking filter and dust protector, which can be simply omitted or replaced by a deep red or complete visible light blocking filter. The Sigma SD14 has an IR/UV blocking filter that can be removed/ installed without tools. The result is a very sensitive digital IR camera.

While it is common to use a filter that blocks almost all visible light, the wavelength sensitivity of a digital camera without internal infrared blocking is such that a variety of artistic results can be obtained with more conventional filtration.

For example, a very dark neutral density filter can be used (such as the Hoya ND400) which passes a very small amount of visible light compared to the near-infrared it allows through. Wider filtration permits an SLR viewfinder to be used and also passes more varied colour information to the sensor without necessarily reducing the Wood effect.

Wider filtration is however likely to reduce other infrared artefacts such as haze penetration and darkened skies. This technique mirrors the methods used by infrared film photographers where black and white infrared film was often used with a deep red filter rather than a visually opaque one.

Several Sony cameras have the so-called Night Shot facility, which physically moves the blocking filter away from the light path, which makes the cameras very sensitive to infrared light.

Soon after its development, this facility was 'restricted' by Sony to make it difficult for people to take photos that saw through clothing. To do this the iris is opened fully and exposure duration is limited to long times of more than 1/30 second or so. It is possible to shoot infrared but neutral density filters must be used to reduce the camera's sensitivity and the long exposure times mean that care must be taken to avoid camera-shake artifacts.

Fuji have produced digital cameras for use in forensic criminology and medicine which have no infrared blocking filter. The first camera, designated the S3 PRO UVIR, also had extended ultraviolet sensitivity (digital sensors are usually less sensitive to UV than to IR).

Optimum UV sensitivity requires special lenses, but ordinary lenses usually work well for IR. In 2007, FujiFilm introduced a new version of this camera, based on the Nikon D200/FujiFilm S5 called the IS Pro, also able to take Nikon lenses. Fuji had earlier introduced a non-SLR infrared camera, the IS-1, a modified version of the FujiFilm FinePix S9100. Unlike the S3 PRO UVIR, the IS-1 does not offer UV sensitivity. FujiFilm restricts the sale of these cameras to professional users with their EULA specifically prohibiting "unethical photographic conduct".

Phase One digital camera backs can be ordered in an infrared modified form.

Remote sensing and thermographic cameras are sensitive to longer wavelengths of infrared. They may be multispectral and use a variety of technologies which may not resemble common camera or filter designs. Cameras sensitive to far infrared including those used in infrared astronomy often require cooling, since all objects at room temperature (including the camera body, optics, and the detector itself) are glowing all the time at these wavelengths. Lower priced thermographic digital cameras extend the spectral range less far into infrared. These cameras are generally used for building inspection or preventative maintenance but can be used for artistic pursuits as well, such as this image of a cup of coffee.

GEOMETRIC CHARACTERISTICS OF AERIAL PHOTOGRAPHS

The Single Vertical Aerial Photograph

The geometry of an aerial photograph is based on the simple, fundamental condition of *collinearity*. By definition, three or more points that lie on the same line are said to be *collinear*. In photogrammetry, a single ray of light is the straight line; three fundamental points must always fall on this straight line: the imaged point on the ground, the focal point of the camera lens, and the image of the point on the film or imaging array of a digital camera.

Picture the bundle of countless rays of light that make up a single aerial photograph or digital frame image at the instant of exposure. The length of each ray, from the focal point of the camera to the imaged point on the ground, is determined by the height of the camera lens above the ground and the elevation of that point on the ground. The length of each ray, from the focal point to the photographic image, is fixed by the focal length of the lens.

Now imagine that the camera focal plane is tilted with respect to the ground, due to the roll, pitch, and yaw of the aircraft. This will affect the length of each light ray in the bundle, and it will also affect the location of image point in the 2-dimensional photograph. If we want to make precise measurements from the photograph and relate these measurements to real world distances, we must know the exact position and angular orientation of the photograph with respect to the ground. Today, we can actually measure position and angular orientation of the camera with respect to the ground with GPS/IMU direct georeferencing technology. But the early pioneers of photogrammetry did not have this advantage. Instead, they developed a mathematical process, based on the collinearity condition, which allowed them to compute the position and orientation of the photograph based on known points on the ground. This geometric relationship between the image and the ground is called *exterior orientation*. It is comprised of six mathematical elements: the x, y, and z position of the camera focal point and the three angles of rotation: omega (roll), phi (pitch), and kappa (yaw), with respect to the ground. The mathematical process of computing the exterior orientation elements from known points on the ground is referred to by the photogrammetric term, *space resection*.

If we were together in a classroom, I could demonstrate the concept of space resection using my desktop and a photograph taken of my desktop from above. I could use pieces of string to represent individual rays of light; each string is of a fixed length based on the distance from the desktop to the camera when the photo was taken. You'll now have to try to picture this demonstration as I describe it in words. If you feel frustrated, imagine yourself as Laussedat trying to figure this out by himself back in the 1800s.

I attach one end of one piece of string to a particular point on my desktop, I attach the other end of the string to the image of that point on the photograph, and I pull the string taut. With that single piece of string, I cannot precisely locate or fix the position of the camera focal point or the orientation of the camera focal plane as it was when the photograph was taken. Now, I choose a second point, adding a second piece of string, and I pull both strings taut. I can't move the photograph around as much as I could with only one piece of string attached, but the photograph can still be rolled and twisted with respect to the desktop. If I identify yet a third point (that does not lie in a straight line with the first two) and attach a third piece of string, I now have a rigid solution; the geometric relationship between the desktop and the photograph is fixed, and I can locate the focal point of the camera in my desktop model. Adding more points adds to the strength of the geometric solution and minimizes the effects of any small errors I might have made, either cutting the string to its proper length or attaching the strings exactly to the points I identified. When we overdetermine a solution by adding additional, redundant measurements, we can make statistical calculations to quantify the precision of our geometric solution.

We don't have time in this course to go into the mathematics of analytical photogrammetry, but hopefully you can get a sense of it as a true measurement science. In fact, photogrammetry has traditionally been taught as a subdiscipline of civil engineering and surveying, rather than geography. Photogrammetry is not just about making neat and useful maps; a key function of the photogrammetrist, as a geospatial professional, is to make authoritative statements about the spatial precision and accuracy of photogrammetric measurements and mapping products. As you'll see in this lesson and the ones to come, you can easily be trained to push buttons in software to produce neat and interesting remote sensing products for use in GIS. It takes a more rigorous education to make quantitative statements about the spatial accuracy of those products. In my opinion, it is as much the duty of the photogrammetrist or GIS professional to make end users aware of the error contained in a data product as it is to give them the product in the first place. Understanding errors and the potential consequences of error is a very important part of the decision-making process. There's also a particular language used to articulate statements about error and accuracy.

The Stereo Pair

Once the exterior orientation of a single vertical aerial photograph is solved, other points identified on the photographic image can be projected as more rays of light, more pieces of string, passing through the focal point of the camera and intersecting the target surface (the ground or my desktop). If the target surface is perfectly flat, then the elevations of the three known points determine a mathematical plane representing the entire surface. It is then possible to precisely locate any other point we can identify in the image on the target surface, merely by projecting a single, straight line. In reality, the target surface is never perfectly flat. We can project the ray of light from the image through the focal plane, but we can't determine the point at which it intersects the target surface unless we know the shape of the surface and the elevation of that point. In the context of our demonstration above, we need to know the exact length of the new piece of string to establish the location of the point of interest on the target surface. This is where the concept of stereoscopic measurement comes into play. I've instructed you to skim through the portion of the Jensen chapter that describes methods of stereoscopic viewing. You won't be held responsible for the details, but I hope you come away somewhat amazed by the fact that our own vision itself works according to these principles. We've been able to build an entire science of measurement around something that is a natural, built-in, characteristic of the physical human being.

The advantage of stereo photography is that we can extract 3-dimensional information from it. Let's return to the example I described above. This time, imagine I have two photographs of my desktop taken from two separate vantage points, and that the individual images actually overlap. One image is taken from over the left side of my desk, and the other is taken from over the right side of my desk. The middle of my desktop can be seen in both photographs. Now, let's assume that we have established the exterior orientation parameters for each of the two photographs; so, we know exactly where they both were at the moment of exposure, relative to the desktop. We now have two bundles of light rays, some of them intersecting in the middle of my desktop. It is a fundamental postulate of geometry that if two lines intersect, their intersection is exactly one point. Voila! Now we can precisely locate any point on the desktop surface, regardless of its shape, in 3-dimensional space. For any given point common to both photographs, we now know the exact length of each of the two pieces of string (one from each photograph) that connect to the imaged point on the ground. In photogrammetry, we call this a *space intersection*. If we have two photographs precisely oriented in space relative to each other, we can always intersect two lines to find the 3-dimensional ground coordinate of any point common to both photographs. The two photographs, oriented relatively to each other, are referred to as a *stereo model*.

By extension, we could have a large block of aerial photographs, overlapping in the direction of flight as well as between adjacent flight lines, all oriented relatively to each other. This block, once constructed, represents all of the intersecting bundles of light rays from all of the overlapping photos. You can imagine that many of the points will be seen in a number of photographs. In fact, with 60% forward overlap, every point in a single flight line is seen 3 times. If the point falls in the 30% sidelap area between two flight lines, it will be seen 6 times; six rays will all intersect at one point. Actually, because some degree of measurement error is unavoidable, the intersection will occur within some sort of sphere, which represents the uncertainty in the projected coordinate of the point in question. As I mentioned earlier, the mathematical equations of photogrammetry allow us to quantify this uncertainty in statistical terms. Your readings will take you into greater depth and detail, but I hope my explanation helps you create a 3-dimensional picture in your mind, making the readings easier to understand.

Aerial Photography Platforms

Radio-controlled Aircraft

Advances in radio controlled models have made it possible for model aircraft to conduct low-altitude aerial photography. This has benefited real-estate advertising, where commercial and residential properties are the photographic subject. Full-size, manned aircraft are prohibited from low flights above populated locations. Small scale model aircraft offer increased photographic access to these previously restricted areas. Miniature vehicles do not replace full size aircraft, as full size aircraft are capable of longer flight times, higher altitudes, and greater equipment payloads. They are, however, useful in any situation in which a full-scale aircraft would be dangerous to operate. Examples would include the inspection of transformers atop power transmission lines and slow, low-level flight over agricultural fields, both of which can be accomplished by a large-scale radio controlled helicopter. Professional-grade, gyroscopically stabilized camera platforms are available for use under such a model; a large model helicopter with a 26cc gasoline engine can hoist a payload of approximately seven kilograms (15 lbs).

Recent (2006) FAA regulations grounding all commercial RC model flights have been upgraded to require formal FAA certification before permission to fly at any altitude in USA.

Because anything capable of being viewed from a public space is considered outside the realm of privacy in the United States, aerial photography may legally document features and occurrences on private property.

Types of Aerial Photographs

Oblique Photographs

Photographs taken at an angle are called *oblique photographs*. If they are taken from a low angle earth surface–aircraft, they are called *low oblique* and photographs taken from a high angle are called *high* or *steep oblique*.

Vertical Photographs

Vertical photographs are taken straight down. They are mainly used in photogrammetry and image interpretation. Pictures that will be used in photogrammetry are traditionally taken with special large format cameras with calibrated and documented geometric properties.

Combinations

Aerial photographs are often combined. Depending on their purpose it can be done in several ways, of which a few are listed below.

- Panoramas can be made by stitching several photographs taken with one hand held camera.

- In pictometry five rigidly mounted cameras provide one vertical and four low oblique pictures that can be used together.
- In some digital cameras for aerial photogrammetry images from several imaging elements, sometimes with separate lenses, are geometrically corrected and combined to one image in the camera.

Orthophotos

Vertical photographs are often used to create orthophotos, photographs which have been geometrically "corrected" so as to be usable as a map. In other words, an orthophoto is a simulation of a photograph taken from an infinite distance, looking straight down from nadir. Perspective must obviously be removed, but variations in terrain should also be corrected for. Multiple geometric transformations are applied to the image, depending on the perspective and terrain corrections required on a particular part of the image.

Orthophotos are commonly used in geographic information systems, such as are used by mapping agencies (e.g. Ordnance Survey) to create maps. Once the images have been aligned, or 'registered', with known real-world coordinates, they can be widely deployed. Large sets of orthophotos, typically derived from multiple sources and divided into "tiles" (each typically 256 x 256 pixels in size), are widely used in online map systems such as Google Maps. OpenStreetMap offers the use of similar orthophotos for deriving new map data. Google Earth overlays orthophotos or satellite imagery onto a digital elevation model to simulate 3D landscapes.

Aerial Video

With advancements in video technology, aerial video is becoming more popular. Orthogonal video is shot from aircraft mapping pipelines, crop fields, and other points of interest. Using GPS, video may be embedded with meta data and later synced with a video mapping program. This 'Spatial Multimedia' is the timely union of digital media including still photography, motion video, stereo, panoramic imagery sets, immersive media constructs, audio, and other data with location and date-time information from the GPS and other location designs.

Aerial videos are emerging Spatial Multimedia which can be used for scene understanding and object tracking. The input video is captured by low flying aerial platforms and typically consists of strong parallax from non-ground-plane structures. The integration of digital video, global positioning systems (GPS) and automated image processing will improve the accuracy and cost-effectiveness of data collection and reduction. Several different aerial platforms are under investigation for the data collection.

History of Aerial Photography

There are many interesting events in the history of aerial photographic interpretation/remote sensing. The First Edition of the *Manual of Remote Sensing* of the American Society of Photogrammetry and Remote Sensing has a good chapter; while the Second Edition of the "Manual" has only one part of Chapter One covering history. Some of the material that follows was taken from a wide variety of sources including the Third Edition of the *Manual of Photogrammetry*; books such as *Deep Black* by William Burrows; and *Air Spy*, by Constance Babington Smith; and many Technical Reports and some newspaper articles. The chronology shows that this technology has:

- matured relatively recently;
- been built upon the inputs of a wide variety of individuals, some of whom they have heard of before;
- been driven by both the military and the commercial marketplace; and
- is continuing a rapid technological advance on a global scale (e.g. SPOT, France; Radar-Sat, Canada; JERS-1, Japan; IRS, India; and all the U.S. commercial satellites).

- Chronological History of Aerial Photography and Remote Sensing
- Important dates in the chronological history of photography, aerial photographic interpretation, and remote sensing:
- Circa 300 BCE-Greece, Aristotle philosophizing at some length about the nature of light, envisions light as a quality and not as an actual substance; as it was thought of by many at the time. He observed that some objects have the potential for transparency but this state is only rendered actual by the presence of light. He then defined light as the act of, or energy of, a transparent body as such.
- 10th Century-Al Hazan of Basra credited with the explanation of the principle of the camera obscura.
- 1666-Sir Isaac Newton, while experimenting with a prism, found that he could disperse light into a spectrum of red, orange, yellow, green, blue, indigo, and violet. Utilizing a second prism, he found that he could recombine the colors into white light.
- 1802-Thomas Young puts forth the basic concepts of the Young-Von Helmholtz theory of color vision: Three separate sets of cones in the retina of the eye, one attuned to red, one to blue, and one to green.
- 1827-Joseph Nicephore Niepce takes the first picture of nature. (Exposure time was 8 hours, emulsion was bitumen of Jedea.)
- 1829-Joseph Nicephore Niepce and Louis M. Daguere signed their partnership agreement (Nicephore Niepce had been working on Heliography, or sun drawing; Daguerre on dioramas, which he constructed with the aid of a camera obscura.)
- 1839-Daguerre announces the invention of Daguerrotype (Niepce had died). Daguerre had discovered that mercury vapors could bring out an image on a silver plate and that sodicum thiosulfate ("hypo") could fix the image and make it permanent.

- 1939-William Henry Fox Talbot describes a system of imaging on silver chloride paper using a fixative solution of sodium chloride. Talbot later found that the latent image could be developed in a solution of gallic acid, and he was the first person to employ a negative/positive process "Calotype" laying the groundwork for modern photography.
- 1830s-Invention of the stereoscope by the Germans. The device was used during the Victorian era for amusement.
- 1855-Scottish physicist James Clark Maxwell, postulates the color additive theory for the production of color photographs.
- 1858-First known aerial photograph is taken from a captive balloon from an altitude of 1,200 feet over Paris by Gaspar Felix Tournachon Nadar.
- 1861-With the help of photographer Thomas Sutton, Maxwell demonstrates his techniques using a bow of multicolored ribbon. (Red filter-sulfo-cyanice of iron, blue filter-ammoniacal sulfate of copper, green filter-copper chloride, a fourth filter of lemon-coloured glass was also used.)
- 1860s-Use of aerial observations from captive balloons in American War. Balloons used to map forest in 1862 not aerial photo though.
- 1870s-Pictures taken from greater heights, 33,000-34,000 feet, from free balloons.
- 1873-Herman Vogel found that by soaking silver halide emulsions (which are naturally sensitive to only blue light) in various dyes, he could extend their sensitivity to longer and longer wavelengths, paving the way for photography in the near infrared.
- 1879-S.P. Langley begins work to find a superior radiation detector.
- 1887-Germans began experiments with photography for forestry.

- 1899-George Eastman produced a nitrocellulose-based film which retained the clarity of the glass plates which had been used to that time.
- 1903-Julius Neubronne patents breast mounted camera for pigeons.
- 1906-Albert Maul takes first aerial photograph using a rocket propelled by compressed air which rose to a height of 2,600 feet and took pictures and then parachuted the camera back to earth.
- 1906-G.R. Lawrence who had been experimenting with cameras for some time (some of which weighed more than 1,000 lbs.) which were hoisted into the air with the aid of balloon-kites and associated controls, takes pictures of San Francisco earthquake and f ire damage from an altitude of some 600 meters. Many people have thought that these photos were taken from airplanes. Lawrence's camera alone weighed more than the Wright Brothers plane and its pilot combined.
- 1909-Wilbur Wright takes first aerial photograph from an airplane of Centrocelli, Italy. WWI produced a boost in the use of aerial photography, but after the war, enthusiasm wanted.
- 1914-Lt. Lawes, British Flying Service, first takes airplane over enemy territory.
- 1915-Cameras especially built for aerial use are being produced. Lt. Col. J.T.C. More Brabazon designed and produced the first practical aerial camera in collaboration with Thornton Pickard Ltd.
- 1918-By this time in WWI, French aerial units were developing and printing as many as 10,000 photographs each night, during periods of intense activity. During the Meuse-Argonne offensive, 56,000 prints of aerial photograpy were made and delivered to A merican Expeditionary Forces in four days.
- 1914-1919-WWI produces boost in the use of aerial photography, but after war interest wanes.

- 1919-Canadian Forestry Mapping Program begins.
- 1919-Hoffman first to sense from an aircraft in thermal IR. First books: Lee 1922; Joerg 1923 (urban); Platt & Johnson 1927 (archaeology).
- 1924-Mannes and Godousky patent the first of their work on multi-layer film which led to the marketing of Kodachrome in 1935.
- 1931-Stevens development of an IR sensitive film (B&W).
- 1934-American Society of Photogrammetry founded. Photogrammetric Engineering is first published. This journal of the American Society of Photogrammetry was later renamed Photogrammetric Engineering and Remote Sensing. The Society is now named the Ame rican Society of Photogrammetry and Remote Sensing.
- 1936-Captain Albert W. Stevens takes the first photograph of the actual curvature of the earth-taken from a free balloon at an altitude of 72,000 feet.
- 1920s-1930s-Interest in the peaceful uses of aerial photography increases (ISDA, USAF, TVA). WWII brought about more sophisticated techniques in API.
- 1941-1945-WWII brings about the development of more sophisticated techniques in aerial photographic interpretation (API). American, British and Germans all produce promising TIR devices.
- 1942-Kodak patents first false color IR sensitive film.
- 1946-First space photographs from V-2 rockets.
- 1950s-Advances in sensor technology move into multi-spectral range.
- 1954-Westinghouse develops first side-looking airborne radar system.
- 1954-U-2 takes first flight.
- 1956-Lu Meuser makes first TIR motion picture employing an AN/AAS-4, a devise for air to ground strip mapping ("...features and vehicles move like an old keystone cops movie.")

- 1960-U-2 is "shot down" over Sverdlovsk, USSR.
- 1960-TIROS 1 launched as first meteorological satellite.
- 1960s-U.S. begins collection of intelligence photography from Earth orbiting satellites, CORONA and KH programs.
- 1962-Zaitor and Tsuprun construct prototype nine lens multispectral camera permitting nine different film-filter combinations. ITEK employs camera to explore the potential value of multispectral photography.
- 1964-SR-71 shown to the press in the Presidential campaign between Goldwater and LBJ.
- Late 1960s-Gemini and Apollo Space photography.
- 1968-Hemphill describes first use of laser for airborne sensing.
- 1972-Launch of the first Earth Resources Technology Satellite (ERTS-1). This system is later renamed Landsat-1. ERTS carries a return beam vidicon (RBV) and a multispectral scanner (MSS).
- 1972-Photography from Sky Lab precursor of manned space station whos first element launch is currently scheduled for 1998.
- 1975-Launch of Landsat 2.
- 1978-Launch of Landsat 3 (March 5).
- 1978-Launch and failure of Seasat. First civil SAR satellite.
- 1978-Launch of Nimbus 7 (Coastal Zone Color Scanner).
- 1978-Launch of NOAA 6 (aka TIROS-N), first satellite to carry the advanced very high resolution radiometer (AVHRR) on board.
- 1981-Launch of SIR-A (Space Imaging Radar-A).
- 1982-Launch of Landsat 4 (Thematic Mapper and MSS).
- 1984-Launch of SIR-B.
- 1984-Launch of Landsat 5.
- 1985-Landsat Commercial contract awarded to EOSAT. Vendor takes over operation of the satellites and rights to Landsat data.

- 1986-Launch of SPOT-1, French Earth Resources Satellite (Systeme Probatoire de la Observation de la Terre.
- 1988-Indian Remote Sensing Satellite (IRS) launched.
- 1990-Launch of SPOT-2.
- 1991-Launch of ERS-1, European Radar Satellite, primarily designed for oceanographic applications.
- 1991-Second Indian Remote Sensing Satellite launched.
- 1992-JERS, Japanese Earth Resources Satellite launched with L-band radar and visible and infrared radiance/ reflectance recording devices on-board.
- 1992-Land Remote Sensing Act of 1992 brings Landsat back under U.S. Government control. EOSAT retains data rights to some Landsat data for up to ten years from acquisition.
- 1993-Launch of SIR-C.
- 1993-Launch of SPOT-3.
- 1994-Landsat 6 fails to achieve orbit.
- 1995-Third Indian Remote Sensing Satellite launched.
- 1995-Canada launches RADARSAT.
- 1995-Early CORONA and KH satellite data are declassified by an Executive Order signed by President Clinton on 23 February. This order authorizes the declassification of intelligence satellite photography acquired in the 1960s.
- 1995-Launch of ERS-2.
- 1995-First indication that a new class of intelligence satellite is being developed appears in the press. The new satellite code name 8x is said to be a major upgrade of the KH-12 spy satellite. The satellite which may weigh as much as twenty tons wil l be able to acquire "intricately detailed images of an area as large as 1,000 square miles of the Earth's surface...with roughly the same precision as existing satellites," according to an article in the September 28 Los Angeles Times. The Time article goes on to say that the current generation of photographic satellites photograph

areas about 10 miles by 10 miles (100 square miles) typically showing details as small as six inches.

- 1997-Proposed launch date of SeaWiFs, replacement for the coastal zone color scanner.
- 1997-Proposed launch of SPOT-4.
- 1998-Proposed first launch of the Earth Observing System's (EOS) AM-1 series on a Polar Orbiting Platform (POP).
- 1998-Proposed launch date for Landsat-7.
- 2000-Proposed launch date of EOS PM-1 series on POP.

Geometry of Aerial Photography

We can define vertical aerial photographs as a photo taken from an aerial platform (either moving or stationary) wherein the camera axis at the moment of exposure is truly vertical. In actuality, vertical airphotos with less than 3* tilt are considered vertical (for most photo interpretation purposes); while those with more than 3* tilt are considered oblique. There are two basic types of oblique aerial photography. These two types are:

1. High angle oblique; and
2. Low angle oblique.

In a high angle oblique, the apparent horizon is shown; while in a low angle oblique the apparent horizon is not shown. Often because of atmosphere haze or other types of obscuration the true horizon of a photo cannot really be seen. However we often can see a horizon in an oblique air photo. This is the apparent horizon.

The basic advantages of vertical air photos are:

1. The scale is essentially constant;
2. Measurements of directions are easier than on oblique photograph. Directions can also be measured more accurately;
3. Within limits a vertical aerial photograph can be used as a map (if grids and marginal data are added); and,
4. Vertical aerial photographs are often easier to interpret than oblique and are better for stereo (there is no masking).

The advantages of an oblique aerial photograph include:

1. Given a constant altitude and camera you can cover a much larger area on a single photo;
2. The view of some objects is more familiar to the interpreter; and,
3. Some objects not visible on vertical photos may be seen on oblique.

(Paine talks about clearance and cloud cover; but that's a tricky one (too cloudy for vertical but maybe enough clearance for an oblique).

Three terms need defining here, they are Principal Point, Nadir and Isocenter. They are defined as follows:

1. Principal Point-The principal point is the point where the perpendicular projected through the center of the lens intersects the photo image.
2. Nadir-The Nadir is the point vertically beneath the camera center at the time of exposure.
3. Isocenter-The point on the photo that falls on a line half-way between the principal point and the Nadir point.

On a true vertical aerial photograph all three of these would be at the same point. There is no such thing as a true vertical aerial photo. All air photos have some degree of tip or tilt.

A quick review.

Vertical Airphotos (0-3* tilt).

3 Photo Centers: Principal Point, Nadir, Isocenter.

These points are important because certain types of displacement and distortion radiate from them. It is the Isocenter of the aerial photo from which tilt displacement radiates. It is Nadir from which topographic displacement radiates.

Perspective and Projection

Now lets talk about perspective and projection.

First lets consider the viewing perspective of a map. On a map objects and features are both planimetrically and geometrically accurate. That is objects are located on the map in exactly the same position relative to each other as they are on the surface of the Earth, except with a change in scale. This is due to the fact that maps use an orthographic projection (i.e. using parallel lines of site) and constant scale to represent features.

Aerial photographs on the other hand are created using a central or perspective projection. Therefore, the relative position and geometry of the objects depicted depends upon the location from which the photo was taken. Now because of this we get certain forms of distortion and displacement in Air Photos.

Distortion and Displacement

There are basically four types of distortions and three types of displacement.

Types of distortion include:

1. Film and Print Shrinkage;
2. Atmospheric refraction of light rays;
3. Image motion; and,
4. Lens distortion.

Types of displacement include:

1. Curvature of the Earth;
2. Tilt; and,
3. Topographic or relief (including object height).

The effects of film shrinkage, atmospheric refraction and the curvature of the Earth are usually negligible in most cases-the exception is precise mapping projects. These types of distortions and displacement will not be discussed here. Image motion will be dealt with further in our lecture on camera systems. That leaves only lens distortion, tilt and topographic displacement to be discussed here. Of these lens distortion is usually the smallest of these. So displacement is typically the largest problem/effect impacting our analyses.

Both distortion and displacement cause changes in the apparent location of objects in photos. The distinction between the types of effects caused lies in the nature of the changes in the photos.

Distortion-Shift in the location of an object that changes the perspective characteristics of the photo.

Displacement-shift in the location of an object in a photo that does not change the perspective characteristics of the photo (The fiducial distance between an object's image and it's true plan position which is caused by change in elevation.)

These types of phenomena are most evident in terrain with high local relief or significant vertical features.

As stated above we will consider here three main types of problems/effects caused by specific types of distortion and displacement. These are the problems/effects associated with:

1. Lens distortion;
2. Tilt Displacement; and,
3. Topographic Displacement.

Lens distortion-Small effects due to the flaws in the optical components (i.e. lens) of camera systems leading to distortions (which are typically more serious at the edges of photos). Car windows/windshields, carnival mirrors are probably the best know examples of this type of effect. These effects are radial from the principal point (making objects appear either closer to, or farther from the principal point than they actually are); and may be corrected using calibration curves.

Tilt Displacement-A tilted photograph presents a slightly oblique view rather than a true vertical record. All photos have some tilt. The perfect gyro stabilization unit, like the perfect lens, has yet to be built. Tilt is caused by the rotation of the platform away from the vertical. This type of displacement typically occurs along the axis of the wings or the flight line. Tilt displacement radiates from the isocenter of the photo and causes objects to be displaced radially towards the isocenter on the upper side of the tilted photo and radially outward on the lower side. If the amount and direction of tilt are known then the photo may be rectified.

Topographic Displacement-This is typically the most serious type of displacement. This displacement radiates outward from Nadir. Topographic displacement is caused by the perspective geometry of the camera and the terrain at varying elevations.

Topographic displacement is not necessarily bad as it allows:

1. Stereo viewing;
2. Height measurement; and,
3. Topographic mapping.

Note: From directly overhead a smokestack looks like a doughnut if there is no shadow.

If the same smokestack is near the edge of the image you can see more of it's side.

Tell the story of trees drown by a reservoir and the displacement pattern on the photo.

The formula for topographic or height displacement on a single photo is:

$$d = r(h)/H = r(h)/(A-E)$$

$$h = d(H)/r = d(A-E)/r$$

Where:

d = Radial displacement (with respect to the Nadir) n the photo at the same scale as the Nadir.

r = Radial photo distance from Nadir (use PP) to the point of displacement (usually the top of the object).

h = Height of the object or difference in elevation (E) between Nadir and displaced point.

H = Flying height above the base of the object (or above the Nadir in some situations).

A close look at the equations involved in the calculations of relief displacement show that some important general relationships are involved. These relationships can be stated as follows:

1. There is no topographic displacement at Nadir. If r is zero, then so is d.
2. Assuming datum elevation to be at Nadir, points above the datum are displaced radially away from Nadir while points below datum are displaced radially towards Nadir.
3. Topographic displacement varies directly with the radial distance from the Nadir to the object. A particular elevation two inches from the Nadir will have half the displacement as that same elevation four inches from the Nadir.
4. Topographic displacement varies directly with the height of an object. A 100 ft. tree would be displaced twice as far as a 50 ft. tree the same distance from Nadir.
5. Topographic displacement varies inversely with the flying height of the base of the object. As a result there is little apparent topographic displacement on space photography.

The reason for small relief displacement from space is that to achieve a given scale a shorter focal length lens requires flying at a lower altitude. The effect of using short focal length lenses is to increase topographic displacement, distortion and the apparent depth of the third dimension (vertical exaggeration) in stereoscopic images).

To get a scale of 1: 20,000 you fly at 10,000 ft. with a six inch focal length lens; but at 20,000 ft. with a 12 inch focal length lens.

Basically, then the most important cause of object displacement on aerial photography is local relief. Remember here that there are times when increased displacement can be a good thing (e.g. for height measurements). So in flat areas you may want to use a short focal length lens to achieve a given scale.

From Space then you can still use extremely long lenses with little displacement.

Orthophotography

Briefly, here there is a growing use of orthophotography today. If you remember back to our discussion of maps at the beginning of the lecture you will remember that orthographic projection depict thing in their true plan position. Basically what happens in the production of orthophotographs is that the original photographs are employ to create a stereo-model which is scanned, by very expensive (today) equipment (orthophotoscope), in very small segments; displacements are corrected and the resulting strips are merged to create a photograph that is essentially a map, actually a planimetrically accurate photo-map.

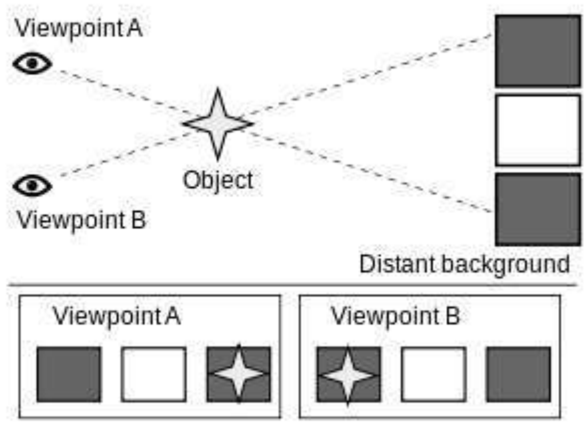
On an orthophoto distances, areas, and directions can all be accurately measured more easily. The U.S. Geological Survey's National Mapping Division is doing all revisions of the 1: 24,000 topographic map updates using digital orthophotography. These ortho photo maps show more detail than the older cartographic product due to the actual image background.

IMAGE DISPLACEMENT, PARALLAX AND STEREOSCOPIC

All aerial images contain geometric distortion. It is impossible to take an aerial image of an area without introducing some sort of error and distortion. Therefore, you should never use aerial imagery that is not been corrected for these geometric distortions. What you should use instead, is an orthographic photo. An orthographic photo is a photo that is been corrected for geometric distortion by a trained photogrammetrist and now shows positions of objects after being projected onto a common plane. This is now equivalent to looking vertically downward onto a scene from an infinite height. Viewing a location from an infinite height is the only way we can improve all geometric distortion. Since we cannot view from such a distance, photogrammetry is employee algorithms, and equations, to emulate the view from an infinite height.

Parallax

A simplified illustration of the parallax of an object against a distant background due to a perspective shift. When viewed from "Viewpoint A", the object appears to be in front of the blue square. When the viewpoint is changed to "Viewpoint B", the object *appears* to have moved in front of the red square.



This animation is an example of parallax. As the viewpoint moves side to side, the objects in the distance appear to move more slowly than the objects close to the camera. In this case, the blue cube in front appears to move faster than the red cube.

Parallax is a displacement or difference in the apparent position of an object viewed along two different lines of sight, and is measured by the angle or semi-angle of inclination between those two lines. Due to foreshortening, nearby objects show a larger parallax than farther objects when observed from different positions, so parallax can be used to determine distances.

To measure large distances, such as the distance of a planet or a star from Earth, astronomers use the principle of parallax. Here, the term *parallax* is the semi-angle of inclination between two sight-lines to the star, as observed when Earth is on opposite sides of the Sun in its orbit.

These distances form the lowest rung of what is called "the cosmic distance ladder", the first in a succession of methods by which astronomers determine the distances to celestial objects, serving as a basis for other distance measurements in astronomy forming the higher rungs of the ladder.

Parallax also affects optical instruments such as rifle scopes, binoculars, microscopes, and twin-lens reflex cameras that view objects from slightly different angles. Many animals, including humans, have two eyes with overlapping visual fields that use parallax to gain depth perception; this process is known as stereopsis. In computer vision the effect is used for computer stereo vision, and there is a device called a parallax rangefinder that uses it to find range, and in some variations also altitude to a target.

A simple everyday example of parallax can be seen in the dashboard of motor vehicles that use a needle-style speedometer gauge. When viewed from directly in front, the speed may show exactly 60; but when viewed from the passenger seat the needle may appear to show a slightly different speed, due to the angle of viewing.

Visual perception

As the eyes of humans and other animals are in different positions on the head, they present different views simultaneously. This is the basis of stereopsis, the process by which the brain exploits the parallax due to the different views from the eye to gain depth perception and estimate distances to objects. Animals also use *motion parallax*, in which the animals (or just the head) move to gain different viewpoints. For example, pigeons (whose eyes do not have overlapping fields of view and thus cannot use stereopsis) bob their heads up and down to see depth.

The motion parallax is exploited also in wiggle stereoscopy, computer graphics which provide depth cues through viewpoint-shifting animation rather than through binocular vision.

Aerial Stereo Photography

Aerial Stereo Photography is any kind of Stereo Photography taken from a flying object (airplane, kite, balloon etc.). Most professional Aerial Stereo Photographs are taken from special planes which are equipped with cameras at the bottom pointing straight down to the earth.

These (large format) photographs are then used to calculate the height of the terrain (and buildings) below through Photogrammetry.

But even amateurs can easily take Aerial Stereo Photographs from any commercial airliner when following some simple rules, set forth in the following paragraphs.

How to take your own Aerial Photographs

Point the camera perpendicular to the direction of flight. During cruise, assume that the craft points where it's going. This is not so during maneuvers: take-off and landing are nose-high. Also not turns and aerobatic maneuvers. These are probably nothing to worry about, but try to keep it simple.

Light aircraft usually cruise at 100-150 feet per second, that's 30-50 meters per second. Guess the distance to your subject and figure the interocular basis you want. 1/ 20 to 1/ 50 of the subject's distance is good to start with. Then decide the number of seconds between pictures. Try taking a quicker triple instead of a slow pair of pictures to give yourself more options later.

Direct sunlight on the aircraft window causes gross flare and reduction of contrast, so either try to have the window removed (works with small, chartered airplanes) or shoot with the sun in your back. The camera gets the shady side of the aircraft. A polarizing filter helps to improve contrast, too. Forget long focal length lenses as there is too much vibration involved. Use a short focal length instead. If you cannot remove the windows of the airplane (as in commercial jets), clean them thoroughly, inside and out (if possible). Bring along cleaning towels anyway.

Now it's time to make a few quick calculations to determine the time between the exposures.

Example 1: A mountain is 2 miles or 10,000 feet away (always remember that you're only guessing!). You are traveling at a speed of 100 ft./sec. 10,000 ft./50 = 200 ft. stereo base; 10,000 ft./20 = 500 ft. stereo base (now you know that you want anywhere between

200 and 500 ft. between the two exposures). 200 ft./100 ft./sec. = 2 seconds; 500 ft./100 ft./sec. = 5 seconds. A stereo pair should therefore be exposed around 2 to 5 seconds apart. If you have a fast winder, it would be best to shoot the following pattern: 0 - 2 - 5 seconds. These three shots give you 3 different stereo pairs: 2 seconds apart (0 - 2), 3 seconds apart (2 - 5) and 5 seconds apart (0 - 5).

Example 2: A Forest Nymph is 500 meters away. You are traveling at a speed of 50 m/sec. So you want anywhere between 500 m/50 = 10 m interocular distance to 500 m/20 = 25 m interocular distance.

Therefore, the time between the two exposures is between 10 m/50 m/sec. = 0.2 seconds and 25 m/50 m/sec. = 0.5 seconds. Therefore, the two images should be taken less than half a second apart. As you probably can't take the two shots that fast, you're probably going to fast, or you're flying too low.

When doing any calculations, always remember the precision of the original data before condemning using coarse arithmetic. A "quick and dirty" calculation is always sufficient!

All these calculations are correct assuming you're shooting perpendicular to your flight path.

If you were pointing your camera out an angle of, say, 30 degrees from the vertical, then you would divide your calculated seconds by the cosine of 30 degrees (0.866) to get the correct result. Usually, there is not enough difference to worry about.

By the same reasoning, if you were 60 degrees off the vertical, then the time delay would be doubled. Those angles should apply to the bottom of the field of view, rather than the center.

Watch out for clouds - if they're the closest objects, they should be the reference for calculations!

And the most important "rule": always try out a range of interocular distances. If you can afford the film and processing, it doesn't hurt to hedge your bets by taking a series of photos, so you can pick out good stereo pairs later - after all, airplane tickets are pretty expensive compared to the cost of film.

Another very rough rule of thumb which is great when you don't have other criteria for judgment: the second shot should be taken when the view shows a barely perceptible difference from that of the first shot. It helps to remember this rule of thumb even when you calculate the spacing more precisely, since it's intuitive and helps in avoiding order-of-magnitude errors.

8

Introduction to Digital Photogrammetry

PHOTOGRAMMETRY

Photogrammetry is the practice of determining the geometric properties of objects from photographic images. Photogrammetry is as old as modern photography and can be dated to the mid-nineteenth century. In the simplest example, the distance between two points that lie on a plane parallel to the ‘. Photogrammetry is used in different fields, such as topographic mapping, architecture, engineering, manufacturing, quality control, police investigation, angy]], as well as by archaeologists to quickly produce plans of large or complex sites and by meteorologists as a way to determine the actual wind speed of a tornado where objective weather data cannot be obtained. It is also used to combine live action with computer generated imagery in movie post-production; *Fight Club* is a good example of the use of photogrammetry in film.

Algorithms for photogrammetry typically express the problem as that of minimizing the sum of the squares of a set of errors. This minimization is known as bundle adjustment and is often performed using the Levenberg–Marquardt algorithm

Photogrammetry uses methods from many disciplines including optics and projective geometry. The data model on the right shows what type of information can go into and come out of photogrammetric methods.. Each of the four main variables can be an *input* or an *output* of a photogrammetric method.

Photogrammetry has been defined by ASPRS as the art, science, and technology of obtaining reliable information about physical objects and the environment through processes of recording, measuring and interpreting photographic images and patterns of recorded radiant electromagnetic energy and other phenomena. Photogrammetric data with dense range data from scanners complement each other. Photogrammetry is more accurate in the x and y direction while range data is generally more accurate in the z direction.

This range data can be supplied by techniques like LiDAR, Laser Scanners (using time of flight, triangulation or interferometry), White-light digitizers and any other technique that scans an area and returns x, y,z coordinates for multiple discrete points (commonly called "point clouds"). Photos can clearly define the edges of buildings when the point cloud footprint can not. It is beneficial to incorporate the advantages of both systems and integrate it to create a better product. Techniques such as adaptive least squares stereo matching are then used to produce a dense array of correspondences which are transformed through a camera model to produce a dense array of x, y, z data which can be used to produce digital terrain model and orthoimage products.

Systems which use these techniques, e.g the ITG system were developed in the 1980s and 1990s but have since been supplanted by LiDAR and Radar based approaches, although these techniques may still be useful deriving elevation models from old aerial photographs and/or satellite images.

This method is commonly employed in collision engineering, especially with automobiles. When litigation for accidents occur and engineers need to determine the exact deformation present in the vehicle, it is common for several years to have passed and the only evidence that remains are crime scene photographs taken by the police. Photogrammetry is used to determine how much the car in question was deformed, which relates to an amount of energy required to produce that deformation. The energy can then be used to determine important information about the crash (like velocity at time of impact).

BASICS OF DIGITAL PHOTOGRAMMETRY

Photogrammetry as a science is among the earliest techniques of remote sensing. The word photogrammetry is the combination of three distinct Greek words 'Photo', 'Gram' and 'metry' which translated in English literally means, light, drawing and measurement respectively. For the laymen, photogrammetry is the technological ability of determining the measurement of any object by means of photography. To better understand the workings or the scope of photogrammetry, it is imperative to know the basic definition of remote sensing described as the process of determining desired information about a situation or an object from a distance; this essentially means without any physical contact with the object.

There are two types digital photogrammetry

For starters, there are 2 distinctive types of digital photogrammetry, 'close-range' and 'aerial'. As the name suggests, close range digital photogrammetry involves the use of a camera mounted on a tripod stand of vehicle or a hand held devise. In this case, the individual images captured by the camera in order to create a 3D picture of the object. This type of photogrammetry technique called 3D Texturing which can be used to create building facades (3D textured building) or any such physical objects those can be viewed from the ground level.

Aerial photogrammetry is done by mounting the camera in an aircraft and subsequently taking photographs of the ground. It is typically done in a straight-down manner. The shutter of the camera is put in to action every few moments as the aerial vehicle moves along the flight path. Previously, film cameras were deployed to capture the images however with the advent of digital technology, digital camera are used in a more efficient manner. These cameras come in very handy when it comes to manipulating the captured images in to something more useful. No matter what type of camera is utilized the captured frames are overlapped with the subsequent chain of frames. These chains of images are then metaphorically 'stitched together' in to a seamless order to make a wholesome picture of the object in interest. Typically, aerial photogrammetry is deployed for mapping terrain, Planimetric data creation, and 3D layer capturing

Photogrammetry for 3D Imagery

A valuable and novel application of photogrammetry is the creation of 3D models. This is sometimes also known as stereo-photogrammetry. This involves the combination of two imagery captured of the same object but from slightly varying angles of having overlap of 60%. Digital photogrammetry helps in producing DEM, DTM , DSM generation, stereo compilation, topographic and Planimetric feature extraction(2D and 3D), Ortho generation and True Ortho generation etc. These models provide the on-looker with 3D pictographic evidence of the object on the ground.

Photogrammetry can also be used in combination with other technologies such as Light detection and Ranging (LiDAR) to create more precise information for city planners, mining experts, geologists, archaeologists, engineers and to anyone else who has a vested interest in visual map of an area.

THE FUTURE OF SPATIAL DATA COLLECTION

PHOTOGRAMMETRY HAS BEEN TRADITIONALLY defined as the process of deriving metric information about an object through photo measurements. A closely related area is photo interpretation, which mainly deals with photographs through human visual analysis. The term "remote sensing" was introduced due to the use of imagery in a wider electromagnetic spectrum as well as the employment of computer analysis techniques.

With the advent of computing and imaging technology, photogrammetry has evolved from analogue to analytical to digital (softcopy) photogrammetry. The term "softcopy photogrammetry" is widely used in the United States for historic reasons, although "digital photogrammetry" is preferred in the international community. The main difference between digital photogrammetry and its predecessors—analogue and analytical—is that it deals with digital imagery directly rather than (analogue) photographs. However, the mathematics of data processing models (e.g., orientation, triangulation, etc.) used in digital photogrammetry has been well established.

Imagery Evolution

The fundamental goal of photogrammetry is to rigorously establish the geometric relationship between an object and an image and derive information about the object strictly from the image. In analogue photogrammetry, optical and mechanical instruments (plotters) are used to establish geometric relationships. In analytical photogrammetry, the geometric modelling is mathematical. Both deal with analogue photographs using expensive photogrammetric plotters.

An analytical plotter from Leica Geosystems GIS & Mapping Division (formerly LH Systems). In digital photogrammetry, imagery of all types, including passive (e.g., optical sensing) or active (e.g., radar imaging), and taken from any platform (e.g., airborne, satellite, close range, etc.) can be processed. Thanks to digital computing technology, the entire photogrammetric production process is digital, and many components have been automated.

Photogrammetry has been used for a variety of applications, ranging from engineering design to natural resource and environmental inventory to hydrographic survey to archaeological mapping. Most topographic maps available today were created using photogrammetry, which is considered the primary approach to GIS base data collection and updating.

Project Requirements

In general, a photogrammetric project involves two stages: 1) acquisition of imagery and its support data (e.g., ground-control information) and 2) processing the imagery to derive image and vector products. The first stage involves several operations such as project design, mission planning, image acquisition, ground control and quality assurance. The second stage involves the use of a digital photogrammetric workstation (DPW) for processing. In the following sections, the processing workflow for the second stage will be addressed, because it's of interest to most GIS practitioners.

A DPW combines computer hardware and software to allow photogrammetric operations to be carried out on digital image data. A list of major commercial DPWs are provided in the accompanying tables on page 36. Typically, a DPW consists of a graphics workstation with, in most but not all cases, a stereo viewing device and a 3-D mouse. For modern DPWs, there's no specific requirement for the host computer. Often a DPW can be built on a high-end desktop PC with at least 256RAM, one or two 19- or 21-inch monitors and a high-performance graphics card.

There are many ways to provide stereo viewing, including a split screen with a simple stereoscope, anaglyph (red/green display), polarization and "CrystalEyes" methods. The choice of device largely depends on operational requirements and cost factors. It also has been realized that many operations in digital photogrammetry don't require stereo viewing, except for 3-D coordinate measurement. Therefore, some DPWs don't provide the stereo viewing devices or make them operational. To maximize mapping efficiency, accuracy and operator comfort, many 3-D control devices have been used, including free-moving hand controllers, hand wheels and foot disks. Two types, the free-hand controller and 3-D mouse, are popular in DPWs.

Most DPWs are based on Microsoft Windows NT or 2000, although most were operated under Unix a few years ago.

Surprisingly, photogrammetry has made a fast transition from Unix or other operating systems to Windows. All the vendor products in the accompanying table support Windows NT or 2000.

Digital Photogrammetry Workflow

A typical workflow in digital photogrammetry. The primary products are digital elevation models (DEMs), ortho-rectified images (or ortho-images) and extracted features (vectors). The secondary products derived from primary products include contour maps (derived from DEMs), image contour maps (ortho-images with contour overlays), image line maps (ortho-images with vector overlays) and 3-D scene models (DEMs with draped images and 3-D features).

Image scanners often are part of the digital photogrammetry environment. Geometric quality and scanning resolution are two major concerns. Typically, 10-20 μm pixel-scanning resolution is required for mapping purposes. Many vendors provide photogrammetric-grade scanners, but they're expensive. Some studies show that the medium- and low-cost desktop scanners could produce reasonable results in terms of geometric quality if appropriately calibrated.

Digital Image Acquisition

The cost of digital imaging sensors decreased dramatically in recent years due to increased market penetration and user acceptability. Many new imaging sensors (e.g., satellite sensors, airborne digital cameras, etc.) have been invented. Airborne digital camera systems are being used in large operational projects and are expected to grow rapidly in the next five years. Commercial remote-sensing systems such as Space Imaging's IKONOS and Digital Globe's Quick Bird satellites further boost the market of image acquisition. Due to high orbital repeatability, stereo and Multispectral capability, and stable and affordable data acquisition costs, digital satellite imagery will have a significant effect on digital photogrammetry.

Orientation and Triangulation

Orientation and triangulation are fundamental photogrammetry operations. Orientation is used to recover the geometric relationship between an object and the image captured. Interior, relative and absolute orientations are some basic modules used to derive 3-D coordinate information of objects from imagery. Automation in interior (inner) and relative orientation is implemented in most DPWs. Triangulation or "called-block adjustment" is used to determine the orientations of all images simultaneously, yielding more accurate and consistent results across the entire mapping area.

Without triangulation, every stereo model would need to be oriented for 3-D coordinate measurements. Triangulation is considered the most important economic factor in photogrammetric mapping. Many DPWs support highly automated triangulation with internally developed modules or third-party solutions such as MATCH-AT developed by INPHO Gmbh.

DEM Generation

One of the most fundamental processes in photogrammetry is to identify and measure corresponding points in two or more overlapped photographs or images. In DPWs, users can attempt to do it automatically—a process called image matching that has been one of hottest research topics in the last 10 years. After finding corresponding points by image matching, a DEM can be generated automatically.

PHOTOGRAMMETRY: RECENT ADVANCEMENTS AND APPLICATIONS

Photogrammetry

Photogrammetry is the practice of determining the geometric properties of objects from photographic images. Photogrammetry is as old as modern photography and can be dated to the mid-nineteenth century. In the simplest example, the distance between two points that lie on a plane parallel to the photographic image plane can be determined by measuring their distance on the image, if the scale (s) of the image is known. This is done by multiplying the measured distance by $1/s$.

A more sophisticated technique, called stereophotogrammetry, involves estimating the three-dimensional coordinates of points on an object. These are determined by measurements made in two or more photographic images taken from different positions. Common points are identified on each image. A line of sight (or ray) can be constructed from the camera location to the point on the object.

It is the intersection of these rays (triangulation) that determines the three-dimensional location of the point. More sophisticated algorithms can exploit other information about the scene that is known *a priori*, for example symmetries, in some cases allowing reconstructions of 3D coordinates from only one camera position.

Photogrammetry is used in different fields, such as topographic mapping, architecture, engineering, manufacturing, quality control, police investigation, and geology, as well as by archaeologists to quickly produce plans of large or complex sites and by meteorologists as a way to determine the actual wind speed of a tornado where objective weather data cannot be obtained. It is also used to combine live action with computer-generated imagery in movie post-production; *Fight Club* is a good example of the use of photogrammetry in film.

Algorithms for photogrammetry typically express the problem as that of minimizing the sum of the squares of a set of errors. This minimization is known as bundle adjustment and is often performed using the Levenberg–Marquardt algorithm.

Photogrammetric Methods

Photogrammetry uses methods from many disciplines, including optics and projective geometry. The data model on the right shows what type of information can go into and come out of photogrammetric methods.

The *3D co-ordinates* define the locations of object points in the 3D space. The *image co-ordinates* define the locations of the object points' images on the film or an electronic imaging device. The *exterior orientation* of a camera defines its location in space and its view direction. The *inner orientation* defines the geometric parameters of the imaging process. This is primarily the focal length of the lens, but can also include the description of lens distortions. Further *additional observations* play an important role: With *scale bars*, basically a known distance of two points in space, or known *fix points*, the connection to the basic measuring units is created.

Each of the four main variables can be an *input* or an *output* of a photogrammetric method.

Photogrammetry has been defined by the American Society for Photogrammetry and Remote Sensing (ASPRS) as the art, science, and technology of obtaining reliable information about physical objects and the environment through processes of recording, measuring and interpreting photographic images and patterns of recorded radiant electromagnetic energy and other phenomena.

Integration

Photogrammetric data with dense range data from scanners complement each other. Photogrammetry is more accurate in the x and y direction while range data is generally more accurate in the z direction. This range data can be supplied by techniques like LiDAR, laser scanners (using time of flight, triangulation or interferometry), white-light digitizers and any other technique that scans an area and returns x, y, z coordinates for multiple discrete points (commonly called "point clouds"). Photos can clearly define the edges of buildings when the point cloud footprint can not. It is beneficial to incorporate the advantages of both systems and integrate them to create a better product.

A 3D visualization can be created by georeferencing the aerial photos and LiDAR data in the same reference frame, orthorectifying the aerial photos, and then draping the orthorectified images on top of the LiDAR grid. It is also possible to create digital terrain models and thus 3D visualisations using pairs (or multiples) of aerial photographs or satellite (e.g. SPOT satellite imagery). Techniques such as adaptive least squares stereo matching are then used to produce a dense array of correspondences which are transformed through a camera model to produce a dense array of x, y, z data which can be used to produce digital terrain model and orthoimage products. Systems which use these techniques, e.g. the ITG system, were developed in the 1980s and 1990s but have since been supplanted by LiDAR and radar-based approaches, although these techniques may still be useful in deriving elevation models from old aerial photographs or satellite images.

Applications

This method is commonly employed in collision engineering, especially with automobiles. When litigation for accidents occurs and engineers need to determine the exact deformation present in the vehicle, it is common for several years to have passed and the only evidence that remains is crime scene photographs taken by the police. Photogrammetry is used to determine how much the car in question was deformed, which relates to the amount of energy required to produce that deformation. The energy can then be used to determine important information about the crash (such as the velocity at time of impact).

Areas of Applications of Close-range Photogrammetry

The ever-expanding areas of application of close-range photogrammetry can be grouped into three major areas: architectural photogrammetry, biomedical and bioengineering photogrammetry (biostereometrics) and industrial photogrammetry. It is noteworthy that the very first measurements ever made by photogrammetry (in the middle of the 19th century) had to do with monuments. It is also a fact that an architect introduced the term "photogrammetry",

Albrecht Meydenbauer, who made his first photogrammetric surveys in 1867. For over century, photogrammetric methods and equipment have continued to evolve. More recently, the field of architectural application of photogrammetry has undergone considerable expansion both in scope and diversity.

Surveys of Historical Monuments

Photogrammetric surveys of historic monuments can be grouped in three major categories: rapid and relatively simple surveys, accurate and complete surveys, and very accurate surveys.

a) Rapid And Relatively Simple Surveys: These are used in preliminary studies for restoration and improvement, in inventory work, and in the study of the history of art. Stereometric cameras and other small format photogrammetric cameras are used extensively, together with "normal case stereo plotters". Plotting is generally at a scale of 1:100. To simplify the operations, inclined photography is taken at standard angles and slope calculators are used.

b) Accurate and Complete Surveys: These are used for systematic documentation of architectural heritage. Plotting scale is generally 1:50, while the details are mapped at 1:20 or 1:10. Large-format metric cameras with long focal lengths are preferred in this type of work in view of the accuracy requirements and the sizes of buildings surveyed. The recently developed wide-angle cameras having local lengths ranging between 100 mm to 150 mm are particularly suitable for this class of photogrammetric surveys.

Accurate surveys are used to document the technical history of the construction of the monument and its evolution as time passes, also to analyse its structural lines and to document its condition and its need for conservation and restoration. This is why one needs high accuracy and precision and as detailed a survey as possible. The use of "first order" stereo plotters is, therefore, essential. Furthermore, normal case photography is often not possible due to the difficult conditions frequently encountered. In some countries, precision photogrammetric surveys have been made for "technical monuments" such as ancient bridges and viaducts of artistic value.

A special case of accurate photogrammetric surveys is the survey of building exteriors (facades). Such surveys are carried out, particularly in central Europe, for the systematic documentation of harmonious architectural groups formed by series of houses in a street or on a square in ancient urban centers in towns and villages. Because of space limitations, facade photography is often taken at an upward inclination (e.g. 30^g or 70^g) or from an elevatable platform on a special truck.

A second special case of accurate photogrammetric survey is the partial detailed survey of particular parts of monuments. Such surveys are conducted in conjunction with restoration and consolidation projects. The highest possible accuracy is needed for these purposes. Depending on the needs, the final outputs of the survey can be in the form of plans, cross-section s, elevations, profiles (for arches), contours (for vaults and cupolas), and/or numerical data giving accurate dimensions between the main elements of the building or distances between these elements.

The photogrammetric surveys conducted in the framework of UNESCO campaigns to salvage prestigious monuments such as About Simbel, Philae, Petra, Borobudur, are good examples of accurate and complete photogrammetric survey.

Very Accurate Photogrammetric Surveys: These are needed for highly refined studies. Accuracy requirement is generally in the order of 1 mm and in some cases 0.1 mm. The study of sculptures in monuments and the assessment of the evolution in the surface of defaced stones (in support of chemical and physical investigations into the "disease of the stone") require this very high accuracy. The principal difficulty in such cases is encountered in photography. Metric cameras permitting short object distances (e.g. by having variable principal distance or through the use of additional lenses) are of great help in this type of work.

Operational Procedures

Procedures for all of the above-discussed types of photogrammetric surveys are well established and documented. Independent stereopairs of photographs are taken either horizontally, vertically or at some inclination using the camera(s) most suitable for the individual project. Base-to-distance ratio is kept rather small (1/5 to 1/15). External controls are kept as simple as possible (such as number of distances and checks on the levelling bubbles of the camera). In case of complex object, however, a network of reference points is necessary. Camera stations are normally located on the ground, on scaffoldings, on nearby buildings, on a hydraulic lift truck or even in helicopters, which are sometimes used to take horizontal photographs of the upper portions of tall buildings.

The photographs are catalogued and stored in "photogrammetric archives" and are plotted only when the need arises. Plotting is mostly done using analogue instruments.

In some photogrammetric surveys, rectifying and assembling a ground of photographs produce photo-plans. This technique is particularly suitable for plane surface of murals, for mosaics, for windows and for facades, particularly when the streets are narrow. In this case, photography is systematically taken at a given inclination and rectified in simple instruments. This approach is appealing both from the technical and economical view.

Orthophotography has recently been experimentally applied in Italy on photography of interiors of cupolas and in Germany and Poland for decorated surfaces with some relief.

Some architectural surveys are made by analytical photogrammetry. In this approach, a certain number of points are accurately determined and then connected by architectural lines. Such a method is not applicable to complex and important monuments because it often involves too many assumptions on the course of the lines to be drawn between points thus giving a theoretical rather than a real representation. On the other hand, analytical methods can be advantageously applied in schematically treating groups of simple constructions, as has been done in Scandinavian countries.

The analytical approach is particularly suitable in studying the structure of monuments and in checking on their stability through the use of digital models. By forming digital models encompassing the monument's fundamental points and the skeleton of its structure, one can study the proportions, define the volumes, compare the form etc. By repeating these operations at intervals of time, one can follow and measure eventual deformations in the building and thus check on its stability. By targeting the points involved in the analytical increases the precision of the observations and the accuracy of the solution.

Both the analogue and the analytical approach lead to numerical data, which is used to determine architectural forms. Using a computer, one can determine the curve of surface that best fits the group of points measured, according to the method of least squares.

Biostereometrics (Biomedical and Bioenginnering Applications of Photogrammetry)

The study of biological form is one of the most engaging subjects in the history of human thought, which is hardly surprising considering the immense variety of living things. As new measurement techniques and experimental strategies have appeared, new fields of inquiry have been launched and more minds have become absorbed with the riddle of biological form. Discovery of the microscope and X-rays prompted the development of microbiology and radiology, respectively. More recently advances in electronics; photo optics, computers and related technologies have helped to expand the frontiers of morphological research. Growing interest in the stereometric analysis of biological form typifies this trend.

Measurements of biological form and function were made from stereophotographs in the middle of the 19th century, shortly after the invention of photography. Why has it taken so long to establish a real place for photogrammetry in the biomedical world? Limited Space does not permit a detailed discussion of this questions: suffice it to say that the problem of bringing photogrammetrist and biomedical specialist together is a bit like trying to unite two tribes who speak different languages and are separated by uncharted territory. In this metaphorical setting, biostereometrics can provide the interpreter-guides needed to negotiate the no-mans land and make more durable connections than those, which have occurred by serendipity alone.

Over the years, contacts between photogrammetrists and biomedical specialists have been quite numerous but most of the contacts involved trying to tie photogrammetry to a particular biomedical speciality. Unfortunately, these efforts generated surprisingly little sustained interaction. Recently, the more wide-ranging approach of systematically relating stereometric analysis to biology and medicine in general has proved to be a more fruitful strategy. If biological structures were regular geometric shapes, there would be no great problem measuring them because simple lengths, breadths, and circumferences would be entirely adequate. But as well all known, organism have irregular three-dimensional components and linear "atomistic" measures such as are produced by tapes and callipers cannot give an unambiguous, comprehensive spatial qualification of a part of an organism as a whole. Biologist and medical specialists are showing renewed interests in the stereometric analysis of biological form. Recent advances in computer technology and a growing range of stereometric sensing techniques have helped to expose the potential of biostereometrics. As a result, the use of photogrammetry is growing in such fields as: aerospace medicine, anthropometry, child growth and development, dentistry, marine biology, neurology, orthodontics, orthopedics, pediatrics, physiology, prosthetics, radiology, and zoology, to mention a few.

The need for biostereometrics stems for from the fact that linear tape and calliper measurements of inherently irregular three-dimensional biological structures are inadequate for many purposes. When stereometric data are used to fill this information gap the potentials for achieving more realistic models and making a more thorough analysis of biological form and function are far reaching.

However, the best tools in the world confer no advantages unless they are used wisely. A petroleum geologist is expected to have the necessary training to select promising sites for oil exploration. Similarly, training in biostereometrics can be helpful in making decisions about "when" and "where" to use photogrammetry in biomedical research and clinical practise.

As the potentials of stereometric analysis in biology and medicine become more widely recognized, the role of biostereometrics in helping to unravel the complexities of organic form and function should continue to grow. Already, several photogrammetrists have chosen careers in biology and medicine and this number is expected to increase over the next few years.

Industrial Photogrammetry

Photogrammetry has been applied in numerous industrial fields and the potentially for further expansion and growth is seemingly limitless. Industrial photogrammetry has been described as "application of photogrammetry in building construction, civil engineering, mining, vehicle and machine construction, metallurgy, ship building and traffic, with their fundamentals and border subjects, including the phases of research, planning, production engineering, manufacture testing, monitoring, repair and reconstruction. Objects measured by photogrammetric techniques may be solid, liquid or gaseous bodies or physical phenomena, whether stationary or moving, that allow of being photographed" by Meyer (1973). The experiences in the fields of architectural photogrammetry and biostereometrics clearly indicate the effectiveness of this strategy. The consistent use of term "industrial photogrammetry" should be instrumental in drawing the attention of photogrammetrist and equipment manufacturers to this fertile field of application, and should be helpful in bringing the capabilities of photogrammetry to the attention of the various industries. Tis way, it is hoped that more and more industrial concerns would make full use of the economical and technical advantage of photogrammetry. Economic benefits of derived from photogrammetric approach has been stated as;

- Measurement time on the object is reduced by %90-%95,
- Saving in manpower,
- Reduce machine and time for blade machining through optimisation of the metal removal rate,
- Reduced material expenditure in the propeller casting manufacture through optimised molds,
- A cut in recycling time for non-ferrous metals,
- Shorter production time for propeller manufacture.

Photogrammetric technique is equally suited for other industries where work-pieces of a complex surface configuration are to be manufactured, which would be very time consuming to measure with conventional measuring tools. A review of selected application in metal working industry has confirmed that photogrammetric techniques can be both practical and economically feasible for industrial measurements inspection tasks. The development of a systematic approach to implementing such application is necessary to investigate the reduction of start-up costs, operating costs and equipment costs.

DIGITAL IMAGE

Analog and Digital Images

An image is a two-dimensional representation of objects in a real scene. Remote sensing images are representations of parts of the earth surface as seen from space. The images may be analog or digital. Aerial photographs are examples of analog images while satellite images acquired using electronic sensors are examples of digital images.

Pixels

A digital image comprises of a two dimensional array of individual picture elements called pixels arranged in columns and rows. Each pixel represents an area on the Earth's surface. A pixel has an intensity value and a location address in the two dimensional image.

The intensity value represents the measured physical quantity such as the solar radiance in a given wavelength band reflected from the ground, emitted infrared radiation or backscattered radar intensity. This value is normally the average value for the whole ground area covered by the pixel.

The intensity of a pixel is digitised and recorded as a digital number. Due to the finite storage capacity, a digital number is stored with a finite number of bits (binary digits). The number of bits determine the radiometric resolution of the image. For example, an 8-bit digital number ranges from 0 to 255 (i.e. 2^8-1), while a 11-bit digital number ranges from 0 to 2047. The detected intensity value needs to be scaled and quantized to fit within this range of value. In a Radiometrically Calibrated image, the actual intensity value can be derived from the pixel digital number.

The address of a pixel is denoted by its row and column coordinates in the two-dimensional image. There is a one-to-one correspondence between the column-row address of a pixel and the geographical coordinates (e.g. Longitude, latitude) of the imaged location. In order to be useful, the exact geographical location of each pixel on the ground must be derivable from its row and column indices, given the imaging geometry and the satellite orbit parameters.

"A Push-Broom" Scanner: This type of imaging system is commonly used in optical remote sensing satellites such as SPOT.

The imaging system has a linear detector array (usually of the CCD type) consisting of a number of detector elements (6000 elements in SPOT HRV). Each detector element projects an "instantaneous field of view (IFOV)" on the ground. The signal recorded by a detector element is proportional to the total radiation collected within its IFOV. At any instant, a row of pixels are formed. As the detector array flies along its track, the row of pixels sweeps along to generate a two-dimensional image.

Multilayer Image

Several types of measurement may be made from the ground area covered by a single pixel. Each type of measurement forms an image which carry some specific information about the area. By "stacking" these images from the same area together, a multilayer image is formed. Each component image is a layer in the multilayer image.

Multilayer images can also be formed by combining images obtained from different sensors, and other subsidiary data. For example, a multilayer image may consist of three layers from a SPOT multispectral image, a layer of ERS synthetic aperture radar image, and perhaps a layer consisting of the digital elevation map of the area being studied.

Multispectral Image

A multispectral image consists of a few image layers, each layer represents an image acquired at a particular wavelength band. For example, the SPOT HRV sensor operating in the multispectral mode detects radiations in three wavelength bands: the green (500-590 nm), red (610-680 nm) and near infrared (790-890 nm) bands. A single SPOT multispectral scene consists of three intensity images in the three wavelength bands. In this case, each pixel of the scene has three intensity values corresponding to the three bands.

A multispectral IKONOS image consists of four bands: Blue, Green, Red and Near Infrared, while a landsat TM multispectral image consists of seven bands: blue, green, red, near-IR bands, two SWIR bands, and a thermal IR band.

Superspectral Image

The more recent satellite sensors are capable of acquiring images at many more wavelength bands. For example, the MODIS sensor on-board the NASA's TERRA satellite consists of 36 spectral bands, covering the wavelength regions ranging from the visible, near infrared, short-wave infrared to the thermal infrared. The bands have narrower bandwidths, enabling the finer spectral characteristics of the targets to be captured by the sensor. The term "superspectral" has been coined to describe such sensors.

Hyperspectral Image

A hyperspectral image consists of about a hundred or more contiguous spectral bands. The characteristic spectrum of the target pixel is acquired in a hyperspectral image. The precise spectral information contained in a hyperspectral image enables better characterisation and identification of targets. Hyperspectral images have potential applications in such fields as precision agriculture (e.g. monitoring the types, health, moisture status and maturity of crops), coastal management (e.g. monitoring of phytoplanktons, pollution, bathymetry changes).

Currently, hyperspectral imagery is not commercially available from satellites. There are experimental satellite-sensors that acquire hyperspectral imagery for scientific investigation (e.g. NASA's Hyperion sensor on-board the EO1 satellite, CHRIS sensor onboard ESA's PRABO satellite).

Spatial Resolution

Spatial resolution refers to the size of the smallest object that can be resolved on the ground. In a digital image, the resolution is limited by the pixel size, i.e. the smallest resolvable object cannot be smaller than the pixel size. The intrinsic resolution of an imaging system is determined primarily by the instantaneous field of view (IFOV) of the sensor, which is a measure of the ground area viewed by a single detector element in a given instant in time. However this intrinsic resolution can often be degraded by other factors which introduce blurring of the image, such as improper focusing, atmospheric scattering and target motion. The pixel size is determined by the sampling distance.

A "High Resolution" image refers to one with a small resolution size. Fine details can be seen in a high resolution image. On the other hand, a "Low Resolution" image is one with a large resolution size, i.e. only coarse features can be observed in the image.

9

Basics of Satellite Remote Sensing

DEFINITION OF SATELLITE REMOTE SENSING

In this CD, you will see many remote sensing images around Asia acquired by earth observation satellites. These remote sensing satellites are equipped with sensors looking down to the earth. They are the "eyes in the sky" constantly observing the earth as they go round in predictable orbits.

Effects of Atmosphere

In satellite remote sensing of the earth, the sensors are looking through a layer of atmosphere separating the sensors from the Earth's surface being observed. Hence, it is essential to understand the effects of atmosphere on the electromagnetic radiation travelling from the Earth to the sensor through the atmosphere. The atmospheric constituents cause wavelength dependent absorption and scattering of radiation. These effects degrade the quality of images. Some of the atmospheric effects can be corrected before the images are subjected to further analysis and interpretation.

A consequence of atmospheric absorption is that certain wavelength bands in the electromagnetic spectrum are strongly absorbed and effectively blocked by the atmosphere. The wavelength regions in the electromagnetic spectrum usable for remote sensing are determined by their ability to penetrate atmosphere. These regions are known as the atmospheric transmission windows. Remote sensing systems are often designed to operate within one or more of the atmospheric windows. These windows exist in the microwave region, some wavelength bands in the infrared, the entire visible region and part of the near ultraviolet regions. Although the atmosphere is practically transparent to x-rays and gamma rays, these radiations are not normally used in remote sensing of the earth.

PRINCIPLE OF REMOTE SENSING SATELLITES

United States Government Remote Sensing Programs

The United States began the current phase of Earth observation from space with the launch of the first Landsat satellite (ERTS-1/ Landsat-1) in 1972. Currently, the United States has ten Earth observing satellites in orbit. Three of these are NOAA satellites (NOAA-J, NOAA-K, and NOAA-L) that comprise the NOAA Polar Operational Environmental System (POES). These NOAA satellites all carry the Advanced Very High Resolution Radiometer (AVHRR) sensor that is used for measuring vegetation densities, crop yields, ocean temperatures, forest fire danger zones, and snow cover.

The other seven are NASA satellites: Landsat 5, Landsat 7, Terra, Tropical Rainfall Measuring Mission (TRMM), Earth Probe—TOMS, Quick Scatterometer (QuikScat) and Earth Observing-1 (EO-1). Landsat 7 is used for general Earth observations including forestry, crop monitoring, land cover, land use, and watersheds.

It carries the Enhanced Thematic Mapper Plus (ETM+), which boasts improved data collection capabilities from previous Landsat missions. Terra differs from Landsat in that it is dedicated to observing process more than land features. Terra carries five different sensors, each having unique applications, ranging from land temperature and snow/glacier cover measurements (ASTER), to cloud cover and radiant energy (CERES), to pollution measurements (MOPITT), to aerosol and smoke plume imaging (MISR), to ocean productivity and temperature ranges (MODIS). The five sensors carried on the TRMM satellite are all committed to record tropical and subtropical atmospheric parameters such as rainfall, lightning, and cloud cover. The TOMS sensor carried on the Earth Probe craft observes rates of ozone depletion, daily UV exposure, UV-absorbing aerosols and data on dust, smoke, and ash in the troposphere. The SeaWinds sensor carried on QuikScat uses specialized radar to measure near-surface wind speed and direction.

International Remote Sensing Programs

International efforts have pioneered the development of active remote sensing satellites. Canadian Space Agency's RADARSAT-1 and the European Space Agency's Remote Sensing satellites (ERS-1 and-2) carry radar sensors that emit and record microwave signals, permitting observations independent of weather or daylight conditions.

France, India, Russia, Japan, and the China-Brazil team all operate successful passive satellite programs. France controls the Systeme Pour l'Observation de la Terre (SPOT), which is comprised of satellites SPOT 1, SPOT 2, and SPOT 4. The payload of the SPOT satellites consists of two high-resolution-visible (HRVIR) sensors that can operate in either panchromatic (SPOT pan) or multispectral (SPOT xs) modes with a resolution of 10-20 meters depending on the mode. SPOT has many applications, including land use, water resources research, coastal monitoring, crop production, and deforestation. SPOT-4 also carries a wide-angle (2000 km) system referred to as VEGETATION that will be used for international crop monitoring. The Indian Space Research Organization (ISRO) currently operates four Earth-observing satellites; the most recently launched (IRS-P4/Ocenasat) focuses on oceanic research. Other Earth-observing systems (EOS) include Russia's Resurs-O1 series, Japan's ADEOS system, and the CBERS satellite that is operated jointly by China and Brazil.

Commercial Satellite Systems

The U.S. government has encouraged the development of independent commercial satellites and many U.S. companies have designed and launched their own satellites. Orbital Imaging Corporation (ORBIMAGE) and Space Imaging, Inc. both have successful satellites in orbit that carry high-resolution sensors. ORBIMAGE operates two satellites. The first, OrbView-1, is designed to monitor atmosphere. The second, OrbView-2 (SeaStar), carries a sensor called SeaWiFS (Sea-viewing Wide Field-of-view Sensor) that was developed in conjunction with NASA. SeaWiFS is designed to monitor ocean temperature and productivity. Space Imaging operates one satellite, IKONOS, which boasts 1-meter resolution capabilities in the panchromatic (black and white) range and 4-meter resolution in the multispectral range. IKONOS has applications ranging from imaging coral reefs to aiding highway planning.

Remote Sensing Applications

The potential applications of these satellite sensors are vast. This section briefly describes some of the possible environmental applications, focusing on environmental enforcement, land use planning, forestry, agriculture, water resources, fisheries, wetlands, watersheds, climate change, and disaster management.

Environmental Enforcement

The U.S. Environmental Protection Agency (EPA) conducts four types of satellite and aerial remote sensing projects to support the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA, also known as the Superfund Act), the Resource Conservation and Recovery Act (RCRA), and in other EPA regulatory programs and investigations. The projects are: (1) emergency response to hazardous material release that requires rapid site assessment; (2) single-date analysis to update old data on the current conditions of the site; (3) intensive site analysis of current and historic images, to obtain an understanding of changing conditions over time; and (4) waste site inventories over large areas to locate possible disposal sites. Images from these projects can standalone or be used in conjunction with topographic maps, digital elevation data, and other features stored in GIS databases.

Further use of remote sensing (both satellite and aerial photography) as a tool in environmental forensics is discussed in a two-part paper by Brilis, et al. The paper outlines the general approach to be followed when planning the use of remote sensing in environmental forensics. The accuracy of locational data and the use of metadata are identified as two critical items to ensure that a final image can withstand veracity issues when used for courtroom presentation.

Recently, interest has developed in using satellites to monitor and enforce multilateral environmental agreements (MEAs), such as the Kyoto Protocol. Remote sensing data may be used in the future to ensure compliance with MEA requirements by both direct enforcement and by more indirect means, such as deterring non compliance through high levels of transparency.

Land Use Planning and Change

Passive sensors, including those on the NOAA-AVHRR, IKONOS, Landsat, and SPOT satellites, are used in a broad range of forest and land use applications. These applications include estimations of primary production, biomass, crop yields, and to chart vegetation type, deforestation, desertification, forest boundaries, forest harvest, soil erosion, and bush or forest fires. Landsat 7's EMT+ sensor is especially useful in studying land use change because its data has been archived since the first Landsat mission in 1972. Passive sensors have also been used to observe and monitor changes associated with storm, flood, and fire damage.

Forestry

Forestry applications for passive remote sensors include tree species surveys, monitoring clear cut operations, planning and observing burn areas, and studying successional forest growth.

The U.S. Forest Service (USFS) relies primarily on the data from Landsats 5 & 7 for forest monitoring because of the low cost and large scene size. Landsat data is particularly applicable to forest change monitoring because data from previous Landsat missions is archived and available for accurate comparison with data from the current Landsat mission.

The USFS also uses SPOT data in conjunction with Landsat data to increase the level of detail in sensitive areas. Active sensors, carried on the RADARSAT and ERS satellites, are capable of making course scale distinctions between cover types such as late successional forests, newly planted forests, clear cut forests, burn areas, agricultural areas, and deserts. Active sensors are valuable tools for monitoring crop regulation compliance, forest clearing, and for taking general inventories of world forest densities.

Agriculture

The United States Department of Agriculture (USDA) is conducting research to determine the potential uses of remote sensing (both aerial and satellite) in the agricultural sector. Promising applications include measuring leaf area indices (LAI—a quantitative indicator of leaf stress), identifying soil properties by their spectral signals, evaluating crop productivity, and providing a valuable data source for crop simulation models. A high-tech type of farming known as "precision agriculture," uses satellite data to characterize specific sections of a field by certain variables (such as water or nutrient levels). Once the characteristics and geographic coordinates of the field section are in a computer, additions such as water, pesticides, and fertilizers can be efficiently controlled in response to the specific needs of each section thereby reducing the amount of pollutants introduced to the environment while producing healthier crops.

Water Resources and Fisheries

SeaWiFS is designed to monitor oceans and track water indicators such as turbidity, sediment load and transport, primary production by marine phytoplankton, algal blooms, chlorophyll content, dissolved oxygen, and pH. Other applications include managing coral reefs, monitoring pollution and oil spills, and characterizing and monitoring short-term and long-term fish habitat. Terra's MODIS and AVHRR sensors record observations of sea surface temperature, which is directly relevant to fisheries due to individual species' temperature requirements for survival and propagation. The sensor may also help predict migration routes. Active sensing technologies are capable of measuring sea level, wave height, surface wind speed, current fronts, eddies, and surface temperature, as well as locating ocean floor features such as trenches and seamounts. Active sensors have also been used to track oil spills, effluent discharges, and algal blooms.

Wetlands and Watersheds

Wetlands monitoring may employ a combination of land-observation and ocean-observation satellites. ETM+ data can be used to delineate wetland areas, make topographical observations, and to detect illegal development. Active systems can provide consistent and accurate observations of dynamic wetland parameters such as tidal and seasonal patterns, climate, hydrology, topography, vegetation, and soil type. Satellite data and images can also be used to delineate the flow of water through watersheds, and can even be used to track pollutants. Furthermore, using algal productivity as an indicator, scientists are able to monitor whether high levels of nutrients pollute areas of a watershed.

Climate Change

In the past decade, various ozone-monitoring sensors have been launched to study global climate cycles. These include the TOMS sensor and many of the sensors on Terra and future EOS satellites. AVHRR data from NOAA's POES satellites is used in conjunction with RADARSAT to monitor the polar ice sheets and iceberg movements. The EOS satellites, beginning with the Terra, were designed specifically for monitoring climate conditions, including the observation of aerosols, cloud cover, fires, ocean productivity, pollution, solar radiation, sea ice, and snow cover.

Disaster Management and Emergency Response

Remote sensing technologies can provide the government with the ability to avoid much of the damage caused by unforeseen natural disasters. While weather satellites have monitored hurricanes and tornados since the 1960s, other satellite sensors, such as ETM+ and MODIS, have potential applications for disaster management and response.

Scientists have used ETM+ data to monitor patterns in floods, droughts, beach erosion, and volcanic activity over time. MODIS and ASTER data can forecast severe weather with a great degree of reliability, potentially saving states millions of dollars in unnecessary evacuation and emergency response.

For forest fire emergencies, TOMS data can identify and monitor the occurrence of forest fires, especially in remote areas, while AVHRR data can create maps denoting fire-susceptible areas. NOAA-POES and NOAA-GOES (Geostationary Operational Environmental Satellite) are used to make weather observations including predicting local weather, tracking weather in real time globally and locally, understanding and predicting hurricanes and other severe weather, studying phenomena such as El Nino, La Nina, the Gulf Stream and other global current patterns, and observing the dynamics between the land temperature, ocean processes, and the atmosphere.

DATA FORMATS FOR DIGITAL SATELLITE IMAGERY

Digital data from the various satellite systems supplied to the user in the form of computer readable tapes or CD-ROM. As no worldwide standard for the storage and transfer of remotely sensed data has been agreed upon, though the CEOS (Committee on Earth Observation Satellites) format is becoming accepted as the standard. Digital remote sensing data are often organised using one of the three common formats used to organise image data.

For an instance an image consisting of four spectral channels, which can be visualised as four superimposed images, with corresponding pixels in one band registering exactly to those in the other bands. These common formats are:

- Band Interleaved by Pixel (BIP)
- Band Interleaved by Line (BIL)
- Band Sequential (BQ).

Digital image analysis is usually conducted using Raster data structures - each image is treated as an array of values. It offers advantages for manipulation of pixel values by image processing system, as it is easy to find and locate pixels and their values.

Disadvantages becomes apparent when one needs to represent the array of pixels as discrete patches or regions, where as Vector data structures uses polygonal patches and their boundaries as fundamental units for analysis and manipulation. Though vector format is not appropriate to for digital analysis of remotely sensed data.

TOPOGRAPHIC MAP

A topographic map is a type of map characterized by large-scale detail and quantitative representation of relief, usually using contour lines in modern mapping, but historically using a variety of methods. Traditional definitions require a topographic map to show both natural and man-made features. A topographic map is typically published as a map series, made up of two or more map sheets that combine to form the whole map. A contour line is a combination of two line segments that connect but do not intersect; these represent elevation on a topographic map.

The Canadian Centre for Topographic Information provides this definition of a topographic map:

A topographic map is a detailed and accurate graphic representation of cultural and natural features on the ground.

Other authors define topographic maps by contrasting them with another type of map; they are distinguished from smaller-scale "chorographic maps" that cover large regions, "planimetric maps" that do not show elevations, and "thematic maps" that focus on specific topics.

However, in the vernacular and day to day world, the representation of relief (contours) is popularly held to define the genre, such that even small-scale maps showing relief are commonly (and erroneously, in the technical sense) called "topographic".

The study or discipline of topography, while interested in relief, is actually a much broader field of study which takes into account all natural and man made features of terrain.

History

Topographic maps are based on topographical surveys. Performed at large scales, these surveys are called topographical in the old sense of topography, showing a variety of elevations and landforms. This is in contrast to older cadastral surveys, which primarily show property and governmental boundaries. The first multi-sheet topographic map series of an entire country, the *Carte géométrique de la France*, was completed in 1789. The Great Trigonometric Survey of India, started by the East India Company in 1802, then taken over by the British Raj after 1857 was notable as a successful effort on a larger scale and for accurately determining heights of Himalayan peaks from viewpoints over one hundred miles distant.

Topographic surveys were prepared by the military to assist in planning for battle and for defensive emplacements (thus the name and history of the United Kingdom's Ordnance Survey). As such, elevation information was of vital importance.

As they evolved, topographic map series became a national resource in modern nations in planning infrastructure and resource exploitation. In the United States, the national map-making function which had been shared by both the Army Corps of Engineers and the Department of the Interior migrated to the newly created

United States Geological Survey in 1879, where it has remained since.

1913 saw the beginning of the International Map of the World initiative endeavoring to map all of Earth's significant land areas at 1:1 million scale on about one thousand sheets covering four degrees latitude by six or more degrees latitude. Excluding boundaries each sheet was 44cm high and (depending on latitude) up to 66 cm wide. Although the project eventually foundered, it left an indexing system that remains in use.

By the 1980s centralized printing of standardized topographic maps began to be eroded by databases of coordinates that could be used on computers by moderately skilled end users to view or print maps with arbitary contents, coverage and scale. For example the Federal government of the United States' *TIGER* initiative compiled interlinked databases of federal, state and local political borders and census enumeration areas, and of roadways, railroads, and water features with support for locating street addresses within street segments.

TIGER was developed in the 1980s and used in the 1990 and subsequent decennial censuses. Digital elevation models (*DEM*) were also compiled, initially from topographic maps and stereographic interpretation of aerial photographs and then from satellite photography and radar data. Since all these were government projects funded with taxes and not classified for national security reasons, the datasets were in the public domain and freely useable without fees or licensing.

TIGER and DEM datasets greatly facilitated Geographic information systems and made the Global Positioning System much useful by providing context around locations given by the technology as coordinates. Initial applications were mostly professionalized forms such as innovative surveying instruments and agency-level GIS systems tended by experts. By the mid 1990s increasingly user-friendly resources such as online mapping in two and three dimensions, integration of GPS with mobile phones and automotive navigation systems appeared. As of 2011 the future of standardized, centrally printed topographical maps is left somewhat in doubt.

Uses

Topographic maps have multiple uses in the present day: any type of geographic planning or large-scale architecture; earth sciences and many other geographic disciplines; mining and other earth-based endeavours; and recreational uses such as hiking or, in particular, orienteering, which uses highly detailed maps in its standard requirements.

SATELLITE ORBIT PROPAGATION

Modeling satellite orbit propagation can be accomplished tram a deterministic model or a statistical model of numerous orbital ephemerae. Because the deterministic approach is so well known in classical physics, we do not seek to repeat the equations and algorithms. While fundamental in nature, these equations would still take some ten pages or so for presentation and more for explanation. The deterministic approach is well grounded but complex (Hoots and Roehrich 1980). One of the earliest works on deterministic modeling of satellite orbits around the Earth was by Brouwer (1959). A brief history of the fundamentals of Brouwer's research and the subsequent modifications by others is presented by Hoots (1981). Tapley et al. (2004) provide a thorough discussion of statistical approaches for modeling satellite orbit propagation. Notable works by others (e.g., Emery et al. 1989) for specific sensors on remote sensing satellites have also been conducted. Those interested in developing their own satellite propagation or sensor models are encouraged to consult these fundamental works.

The orbital propagation solution and sensor viewing geometry developed for the Remote Sensing Hazard Guidance System (Hodgson and Kar 2008) was used in this research. We have built our model based on these works but modified it for efficiency and accuracy. The age of the satellite ephemerae (i.e., the location and trajectory of a satellite observed at one moment) is one of the major factors influencing satellite orbit propagation accuracy. Predictions of near-future satellite locations (e.g., hours later) are trivial and very accurate, while predictions for months ahead require greater complexity and numerous additional factors. Our model was based on other demanding applications but was used in a very ideal context here. For this research, we used orbital ephemerae that were collected nightly and thus are less than 24 hours old. Our orbital predictions for day 'n' used the ephemerae from day 'n'-1 (i.e., the preceding day). By using daily ephemerae for predictions within the next 24 hours at the most, we also obviate the major sources of error in satellite orbit prediction, such as satellite maneuvers.

The first step in the orbital propagation process is to obtain the orbits of all the satellites of interest. The orbital ephemerides are obtained on a nightly basis for all remote sensing satellites. The satellite orbits are converted into a Cartesian Earth centered state vector at the particular epoch of interest. The Cartesian state vectors are used to compute classical orbit elements for each of the satellites. These classical orbit elements may be used to propagate the satellite trajectory forward or backward in time from the epoch specified. In addition, the time and state vector for a particular satellite may be determined using angles from the satellite location at the epoch of interest. For example, the state of a satellite when it reaches its northern most position may be determined from the angle measured from the initial position to the point in the orbit where the true anomaly plus argument of periapsis is 90 degrees. Using these classical orbit elements, the parts of the orbit where the spacecraft is on the dark side of the Earth or too far north to be imaged may be quickly eliminated.

The search algorithm then finds the point on the orbit where the satellite is closest to the landfall point defined on the surface. At this point, the range rate is zero and transitioning from negative to positive. These geometrical parameters may be easily computed from the landfall point location and state vectors of the satellite obtained from orbit element propagation. Various geometrical parameters, such as sun elevation angle or range, may be examined to determine if they satisfy the constraints specified. If they are out of tolerance, the satellite will not have an opportunity to collect imagery at this location.

Once the minimum-range point is found, as determined from conic orbit elements, a new initial epoch is computed. The ephemerae database is interrogated for a new satellite state vector for the date/time calculated using the conic orbit element. The entire search algorithm is then repeated using precision satellite state vectors at the exact time of overflight. This insures that the most accurate results are obtained, and conic orbit propagation is only used for relatively short distances.

SATELLITE VIEWING OPPORTUNITIES

To empirically determine the likelihood of future collection opportunities, the spatial and temporal distribution of historic landfall points were assumed to occur in the year 2005. The year 2005 is arbitrary but was convenient for this research as we had a complete daily collection of the satellite orbital ephemerae needed to predict the orbital tracks of each satellite and, subsequently, the satellite viewing opportunities. As noted earlier, using daily orbital ephemerae allows for very high spatial/temporal accuracy. The use of any other year may change the collection opportunities for a single hurricane event but not for the set of landfall events, unless the orbits of the satellites were temporally correlated.

For each day and geographic landfall location of a hurricane, the viewing opportunity for each of three satellites (i.e., Ikonos-2, Quickbird-2, and Orbview-3) was modelled. Determining if the geographic location can be imaged is solved by first propagating the satellite vehicle orbit during the days following the hurricane landfall, and modeling the viewing geometry for the sensors onboard the satellite vehicle. Most satellites with high spatial resolution sensors have a similar orbital period (90-100 min) resulting in similar repeat intervals on the ground (satellite image providers routinely advertise a 2-3 day repeat visit). Imaging opportunities vary among satellite sensors with variations in altitude and off-nadir viewing. The OSA sensor on Ikonos-2 and Orbview-3 can point 45-degrees off-nadir as compared to the 28-degrees of the BGIS200 on Quickbird-2. For most mapping applications (except for stereoviewing applications), off-nadir viewing angles of less than 26 degrees are desirable. At greater angles, the ground surface obscured by the sides of land cover features (buildings, vegetation, etc.) is problematic. The potential swath coverage depicts the possible image collection opportunities; however, it does not represent the area-field-of-view, also called swath width, which would be imaged by most high spatial resolution sensors. The swath width is always smaller (e.g., 16.5-km in width for Quickbird-2 at nadir) than the potential swath coverage.

The likelihood of satellite imagery collection was examined under several scenarios. First, we assumed a single image provider, and a maximum viewing angle of 26 degrees off-nadir was permitted. Historically, remote sensors in a hurricane disaster context have sought imagery from near-nadir views rather than oblique imagery. The 26-degree criterion has been used informally in emergency operations centres. Second, any of the three satellites were allowed but still imposing a 26-degree off-nadir limit. Finally, collection opportunities from 1) single satellites and 9) all three satellites were examined with the maximum viewing angle allowed by the sensors.

Restricted Viewing Angles

If an emergency response agency relied on a single satellite provider to collect imagery over the hurricane landfall area, the likelihood of obtaining imagery within the first 24 hours ranges from 16 to 18 percent. The likelihood increases to 37 to 42 percent during the first 48 hours and from 53 to 67 percent in the first three days. Even with the restriction of 26 degrees maximum viewing angle, the Ikonos-2 satellite-sensor combination has a higher likelihood of collection as its altitude is higher (i.e., 679 km versus-450 km for the other two).

A higher altitude allows the sensor to point at farther planimetric distances with the same viewing angle. Relying on one satellite-sensor combination with this 26 degree off-nadir viewing limit results in an image collection during the first three days of, at best, only 2 out of 3 (-67 percent) hurricane disasters. However, if satellite imagery from any of the three satellites were permitted, then the likelihood increases to 38 percent, 74 percent, and 91 percent within the first 24, 48, and 72 hours after hurricane landfall.

Maximum Viewing Angles

Not surprisingly, when more extreme viewing angles are allowed (e.g., when the agency must have some imagery regardless of whether portions of the landscape are obscured), the likelihood of collecting imagery increases. Ikonos-2 and Orbview-3 can point off-nadir at angles of 45 degrees. The likelihood of collection in the first 24 hours increases to 39 percent using either of theses satellites. The likelihood increases to 97 percent for either Ikonos-2 or Orbview-3 within the first 72 hours after landfall. By relying on all three satellite image providers, the likelihood of collecting in the first 24 hours is 61 percent.

This research estimated the likelihood of using one or more (up to three) satellite imagery sensors for collecting post-hurricane event imagery in the first 24, 48, and 72 hours of landfall. Relying on a single satellite provider would result in only a 39 percent chance of collecting in the first 24 hours, if using maximum off-nadir viewing angles. To improve the collection likelihoods for post-event imagery, multiple satellite providers must be relied upon. In the concept of operations (CONOPS) model, this would inevitably require pre-existing contracts (pre-existing to enable a rapid agreement to collect and pay for the collection) with multiple satellite image providers. Since the completion of this research, Orbview-3 satellite imaging sensor has malfunctioned and is considered inoperable. However, GeoEye-1, with a.41 cm spatial resolution and a 681-km altitude and off-nadir viewing angle of up to 61 degrees is operational. Using the GeoEye-1 satellite in lieu of Orbview-3 would result in likelihood estimates slightly better than Ikonos-2. Other high spatial resolution satellite-sensors, such as Worldview-1, have recently been launched, providing viewing opportunities from at least four different U.S. commercially controlled satellites.

A logical planning scenario might be to use multiple satellite-sensor contracts and stratify the potential damage area into image collection priorities. An image collection with a near-nadir (low off-nadir angles) vantage might be used for the high-priority areas, while greater off-nadir angle opportunities are used for low priority areas.2 Focusing on large-impact areas would further increase the likelihood of collecting imagery across a disaster area.

A final probing scenario for response agencies might be to use both aerial and satellite imaging platforms in a well defined and coordinated approach. By quickly modeling (e.g., GIS-based without imagery the likely damage area, the collection area could be spatially defined, stratified, and prioritized. Available airborne and satellite-sensors could be utilized in a stratified (both temporally and spatially) approach to collect imagery as part of the post-event response.

There is a revolution underway in Caddo archaeology, and its name is remote sensing. The use of remote sensing technologies on Caddo archaeological sites-in conjunction with archaeological excavations-is producing, and will continue to produce, unprecedented characterizations of the internal spatial structure and organization of Caddo villages and mound centres. We illustrate the importance of remote sensing for the present and future of Caddo archaeology by focusing on the remote sensing and archaeological findings from one Caddo village on the Red River in Bowie County, Texas. The Hill Farm site is part of the larger Hatchel site village and was occupied by the Nasoni Caddo between ca. A.D. 1600 and 1700.

The Nasoni Caddo village was visited by Europeans in 1687, 1690, and 1691, most famously by the Spanish expedition led by Don Domingo Teran de los Rios. On this expedition (Hatcher 1932, 1999), a detailed map of this village was drawn that showed a temple mound and a number of village compounds (with house structures, granaries, and outdoor ramadas or arbors). A convergence of large-scale remote sensing (Grealy and Conyers 2006), the smaller-scale remote sensing, geoarchaeological findings (Guccione and Hays 2006), and the archaeological discoveries at the Hill Farm site and other places at the Hatchel site (Perttula 2005) have led us to the specific identification of the Hill Farm site with two of the village compounds depicted on the 1691 Teran map. This in turn opens the door to the possibility that continued archaeological and remote investigations can, in effect, locate and ground truth the 1691 Teran map and different parts of the Nasoni Caddo village.

Remote Sensing in Caddo Archaeology

Archaeogeophysical prospecting using magnetometry, electromagnetic conductivity, electrical resistivity, ground-penetrating radar, and magnetic susceptibility have all been shown to be useful geophysical techniques for locating buried architectural remains on prehistoric Caddo mound sites, such as the George C. Davis site and Horace Cabe Mounds (Walker and Perttula 2007) in northeastern Texas, and at various sites on the Grandview project in southwestern Arkansas. These techniques measure certain physical properties of the soil, such as magnetic conditions, electromagnetic conductivity, or electrical resistivity. Given the proper soil conditions, archaeological features such as fire hearths, post holes, and storage pits differ from the surrounding soil matrix, and can be recognized by distinctive spatial signatures in geophysical data sets from archaeological sites.

At the George C. Davis site, geophysical surveys, using a portable cesium magnetometer, on large blocks of land have proved successful in locating both large and small architectural features, particularly numerous Caddo structures. This work is contributing new information on the overall layout of structures around the mounds and across this large prehistoric site. In some settings, the geophysical survey has successfully located archaeological features in areas that previous subsurface excavations failed to locate (Walker et al. 2003). Shaded relief maps of the magnetic data have also identified a number of unique architectural forms at the site, such as the so-called Button Houses (Creel et al. 2005), that can be compared to the architectural plans of the more than 50 previously excavated structures at the George C. Davis site.

These geophysical examples of Caddo structures are 10-15 m diameter circular structures with four internal magnetic returns that are spatially patterned on the interior of the structure to encircle a central hearth; these no doubt represent large post hole features (support posts) or interior pits. The magnetometer data from the Mound B area at George C. Davis are providing new information about the construction phases of the mound. These data have detected a trench of some kind or a small mound of earth that apparently served to anchor the mound to a specific place on the ground. The remote sensing work has also identified a second ramp on the southern flanks of the mound (Schultz et al. 2004).

On the Grandview project, the Arkansas Archeological Survey used a variety of geophysical surveying equipment to measure magnetic susceptibility, gradiometry, electrical resistance, and electromagnetic conductivity of key areas at Caddo mound and nonmound settlements, and achieved impressive results. The geophysical surveying has detected the locations of burned Caddo houses, fireplaces or hearths within the houses, burial pits, nonburial pits of various sizes and shapes, stockade lines around the Tom Jones mound (3HE40) (Lockhart 2007), and possible paths across the sites.

Ground-truthing excavations of several of the subsurface anomalies identified in the Grandview geophysical survey have been successful in identifying archaeological features, including burned houses with "heaps of fired daub up to a meter deep in places, pieces of carbonized logs, and a floor laden with ash and domestic debris" (Schambach 2001:10). Other excavations exposed clay-lined hearths, and lines of post holes, while excavations placed over "reasonably clear images [from the remote sensing] of house remains netted... the archaeological remains of four houses" (Schambach 2001:10).

The Archaeology of the Hill Farm Site within the Larger Hatchel Site Village

Before discussing the remote sensing findings and our interpretation of them, we review what is known about the archaeology of the Hill Farm site. The site is only one part of the much larger Hatchel site village occupied by the Nasoni Caddo from the eleventh to the eighteenth century (Perttula 2005). The village was apparently comprised of a number of farmstead compounds spread out over many acres, with domestic structures, granaries, arbors, and ramadas. The farmstead compounds may have been separated by compound fences or dividers of trees and bushes. The Hill Farm site has two areas, Areas A and B, with substantial archaeological deposits in both.

The first archaeological investigations at the Hill Farm part of the Hatchel village was completed by the Works Progress Administration (WPA) and the University of Texas at Austin in 1939 (TARL n.d.a). The work was on a small Caddo mound, 21 m in diameter and 0.9 m high, which is 915 m southeast of the main platform mound at the Hatchel site and 610 m southeast of WPA excavations in Village Plot 1 (Perttula 2005). The mound (Mound 2) is in Area A at the Hill Farm site; its exact location is currently unknown. Three other small mounds (0.6-0.9 m in height) were nearby. Just east of Mound 2 was a borrow pit of unspecified size (TARL n.d.a).

The WPA excavations in Mound 2 were in two blocks at the northern and southern parts of the mound. The northern strip covered 150 ft^2 and the southern excavation was 900 ft^2 in size.

These excavations disclosed that the mound was constructed in one episode of a brown clay fill that was 0.9 m thick. It was probably constructed to bury a structure or important public building, although the details of this building are not known. The mound fill was placed atop a 25 cm thick habitation and midden deposit of Late Caddo Texarkana phase age (ca. A.D. 1400-1700). No structural features or post holes were recorded in these submound deposits.

A single extended burial (Burial 1) was excavated in the southern Mound 2 strip. This adult burial originated at the top of the habitation/midden deposits-prior to mound construction-and the grave pit reached to 1.37 m below surface (bs). No funerary offerings were placed with this Caddo individual. There were seven sherds, including a vessel lug, in the burial pit fill. Only 324 artifacts were recovered in the Mound 2 WPA excavations (none of the archaeological deposits were screened during the WPA work). This included 212 ceramic vessel sherds, one celt, 25 pieces of lithic debris, 71 animal bones, and 15 pieces of freshwater mussel shell.

In the 1950s, according to records on file at the Texas Archeological Research Laboratory, University of Texas, two avocational archaeologists (R. King Harris, from Dallas, Texas, and Pete Miroir, from Texarkana, Texas) conducted excavations in what appears to have been Area B at the Hill Farm site. They apparently encountered several Caddo burials and very large post holes from a large structure. No more details on this work are available, unfortunately.

More recently, Gilmore and McCormick reported that an unknown number of Caddo burials had been looted at the Hill Farm site in the 1970s. Their location within the site is also unknown. In 2005, renewed archaeological investigations at the Hill Farm site were completed as part of the Bowie County Levee Realignment project sponsored by the Tulsa District of the U.S. Army Corps of Engineers. During this work (Sundermeyer et al. 2006), 78 shovel tests were excavated in Area A (27 of which contained prehistoric archaeological materials), with another 127 shovel tests (51 contained prehistoric archaeological remains) in Area B.

Additionally, a magnetometer survey was completed in both Area A (6,000 m^2, ca. 1.5 acres) and Area B (3,600 m^2, 0.9 acres) by Walker and Schultz (2006). It is the findings from both the shovel testing and the remote sensing that we discuss in the remainder of this chapter.

The Hill Farm site contains exceptionally well-preserved and buried Late Caddo Texarkana phase archaeological deposits. The archaeological materials are incorporated in Red River alluvial sediments of the abandoned and recently defined Cemetery Meander (Guccione and Hays 2006). These deposits apparently began to accumulate sometime after ca. A.D. 1240.

The site lies on sandy loam point bar alluvial deposits adjacent to an ancient abandoned channel of the Red River. The archaeological deposits are marked by numerous structures of varying sizes, pit features, midden deposits, at least one constructed earthen mound (Mound 2 at Hatchel) and as many as three other small mounds, a borrow pit, and one extended burial.

The recent shovel testing reported by Sundermeyer et al. (2006) identified two distinct areas (A and B) at the Hill Farm site. Area B, the northern area, covers approximately 13,000 m^2 (3.2 acres), while the southern part of the site (Area A) is about 11,000 m^2 (2.7 acres) in size.

The two different site areas are about 160 m apart. Archaeological materials occur to depths of 80100 cm bs. The vast majority of these remains are buried, coming primarily from 20-80 cm bs.

There are two concentrations of ceramic vessel sherds in Area A, with one large and heavy-density cluster covering ca. 2,400 m^2 in the centre of this village locality. A much smaller and lighter scatter of sherds is ca. 15-30 m south of the principal sherd concentration. In Area B, there are 4-5 distinct concentrations of ceramic vessel sherds. These concentrations are 15-40 m apart from each other within this village area, and most of them are situated in the southern part of Area B.

PASSIVE REMOTE SENSING

Passive systems collect data from energy that is reflected or radiated off the Earth's surface and atmosphere. A typical image derived from an infrared passive sensor consists of small equal areas referred to as pixels (7) arranged in regular rows and columns. Each pixel has a numerical value called a digital number (DN) that records the intensity of electromagnetic energy measured for the area of ground represented by the pixel.

The DN range from 0 to some higher number on a gray-scale. Each pixel is also given x and y coordinates to place it. The image can therefore be described in strictly numeric terms on a three-coordinate system with x and y locating the Pixel and z giving the DN displayed as a gray scale intensity value.

Passive sensors are described in terms of their spatial, spectral, and temporal resolutions.

The spatial resolution of a sensor is the smallest area that is recorded as a separate unit (pixel). For instance, one-meter spatial resolution means that one pixel of a digital image represents an area on the Earth's surface measuring one meter in length by one meter in width. Spectral resolution refers to the number and dimension of bands (or wavelengths) of the electromagnetic spectrum that a sensor records. The higher the number of bands, the greater the sensor's ability to distinguish between objects. Temporal resolution, also known as repeat time, is the frequency with which a sensor passes over the same area.

Active Remote Sensing

Active remote sensing devices, on the other hand, emit high-energy electromagnetic radiation and record the relative amount and pattern of the energy that is reflected back. Many of these devices operate at wavelengths that not only penetrate cloud cover, but also vegetative cover and soil surfaces. The tradeoff for greater imaging capabilities, however, is increased complexity in data interpretation, as compared to passive sensor data interpretation.

Data Processing

After the satellite records the data, it is transmitted to a ground station for calibration and storage. The data may undergo various levels of processing before it is made available to the user. These levels range from simply correcting for transmission errors to performing advanced correction and analysis with model algorithms, depending on the needs of the scientists or user.

Once the data has undergone initial processing techniques, users may apply it for various purposes, from the simple production of an enhanced image to the more complex creation of image maps, thematic maps, and spatial databases. The data may also be used to develop statistical observations and graphs of the observed phenomena. To create maps and spatial databases, the initial data must be combined with other spatial data. An effective method to analyse the remote sensing data with reference to other spatial data is in a geographic information system (GIS).

Remote Sensing Data Integration with Geographic Information Systems (GIS)

Geographic information systems (GIS) are defined as computer systems capable of assembling, storing, manipulating, and displaying geographically referenced information (i.e. data points identified with respect to their location). GIS store information about the world as a collection of thematic layers that can be linked together by geography.

Remote sensing data applications and GIS have an established history of interdependency. GIS provides a format to distribute remote sensing data and to derive useable information from the data. Remotely-sensed data is also a critical means to create base GIS maps and update many data layers in the GIS. The integration of remotely-sensed data and GIS is particularly attractive because 1) the conversion of remotely-sensed raster-format data to GIS vector-format data is inexpensive and 2) remote sensing data offers a cost-effective way to visualize large geographic areas in a digital format.

There are two defining features of all GIS: the ability to overlay spatial data and the ability to change as new data becomes available. The first key feature of GIS programs is the capability to overlay multiple sets of databases into a map format that graphically explains the relationships between the data. Spatial data (points, boundaries, and lines) comprise the base of the map and can be supplemented with tabular data and image data (such as that from satellites). This powerful and versatile concept has proven invaluable for solving many real-world problems, from recording details of land use planning applications to modeling global atmospheric circulation cycles. The second key feature of GIS is their status as "dynamic maps" that can be updated and altered as needed. These maps may also be manipulated to perform scientific analyses and to create models of different environments.

In a simplistic example of GIS application, a map of city streets could be combined with latitude/longitude-referenced traffic flow data to create a map that reveals areas of frequent accident occurrence, potential detour routes, and even alternatives to improve traffic routing and alleviate rush hour stress. The same base map also may be reused to show, for example, changes in traffic patterns across time.

OPTICAL AND INFRARED REMOTE SENSING

In Optical Remote Sensing, optical sensors detect solar radiation reflected or scattered from the earth, forming images resembling photographs taken by a camera high up in space. The wavelength region usually extends from the visible and near infrared (commonly abbreviated as VNIR) to the short-wave infrared (SWIR).

Different materials such as water, soil, vegetation, buildings and roads reflect visible and infrared light in different ways. They have different colours and brightness when seen under the sun.

The interpretation of optical images require the knowledge of the spectral reflectance signatures of the various materials (natural or man-made) covering the surface of the earth.

There are also infrared sensors measuring the thermal infrared radiation emitted from the earth, from which the land or sea surface temperature can be derived.

Microwave Remote Sensing

There are some remote sensing satellites which carry passive or active microwave sensors. The active sensors emit pulses of microwave radiation to illuminate the areas to be imaged. Images of the earth surface are formed by measuring the microwave energy scattered by the ground or sea back to the sensors. These satellites carry their own "flashlight" emitting microwaves to illuminate their targets. The images can thus be acquired day and night. Microwaves have an additional advantage as they can penetrate clouds. Images can be acquired even when there are clouds covering the earth surface.

A microwave imaging system which can produce high resolution image of the Earth is the synthetic aperture radar (SAR). The intensity in a SAR image depends on the amount of microwave backscattered by the target and received by the SAR antenna. Since the physical mechanisms responsible for this backscatter is different for microwave, compared to visible/infrared radiation, the interpretation of SAR images requires the knowledge of how microwaves interact with the targets.

Remote Sensing Images

Remote sensing images are normally in the form of digital images. In order to extract useful information from the images, image processing techniques may be employed to enhance the image to help visual interpretation, and to correct or restore the image if the image has been subjected to geometric distortion, blurring or degradation by other factors. There are many image analysis techniques available and the methods used depend on the requirements of the specific problem concerned. In many cases, image segmentation and classification algorithms are used to delineate different areas in an image into thematic classes. The resulting product is a thematic map of the study area. This thematic map can be combined with other databases of the test area for further analysis and utilization.

Visual System

Passive Remote Sensing: The eyes passively senses the radiation reflected or emitted from the object. The sensing system depends on an external source of illumination.

The human visual system is an example of a remote sensing system in the general sense. The sensors in this example are the two types of photosensitive cells, known as the cones and the rods, at the retina of the eyes.

The cones are responsible for colour vision. There are three types of cones, each being sensitive to one of the red, green, and blue regions of the visible spectrum.

Thus, it is not coincidental that the modern computer display monitors make use of the same three primary colours to generate a multitude of colours for displaying colour images. The cones are insensitive under low light illumination condition, when their jobs are taken over by the rods. The rods are sensitive only to the total light intensity. Hence, everything appears in shades of grey when there is insufficient light.

As the objects/events being observed are located far away from the eyes, the information needs a carrier to travel from the object to the eyes. In this case, the information carrier is the visible light, a part of the electromagnetic spectrum.

The objects reflect/scatter the ambient light falling onto them. Part of the scattered light is intercepted by the eyes, forming an image on the retina after passing through the optical system of the eyes. The signals generated at the retina are carried via the nerve fibres to the brain, the central processing unit (CPU) of the visual system. These signals are processed and interpreted at the brain, with the aid of previous experiences. When operating in this mode, the visual system is an example of a "Passive Remote Sensing" system which depends on an external source of energy to operate. We all know that this system won't work in darkness. However, we can still see at night if we provide our own source of illumination by carrying a flashlight and shining the beam towards the object we want to observe. In this case, we are performing "Active Remote Sensing", by supplying our own source of energy for illuminating the objects.

The Planet Earth

The planet Earth is the third planet in the solar system located at a mean distance of about 1.50 x 10^8 km from the sun, with a mass of 5.97 x 10^{24} kg. Descriptions of the shape of the earth have evolved from the flat-earth model, spherical model to the currently accepted ellipsoidal model derived from accurate ground surveying and satellite measurements. A number of reference ellipsoids have been defined for use in identifying the three dimensional coordinates (*i.e.* position in space) of a point on or above the earth surface for the purpose of surveying, mapping and navigation. The reference ellipsoid in the World Geodetic System 1984 (WGS-84) commonly used in satellite Global Positioning System (GPS) has the following parameters:

- Equatorial Radius = 6378.1370 km
- Polar Radius = 6356.7523 km

The earth's crust is the outermost layer of the earth's land surface. About 29.1% of the earth's crust area is above sea level. The rest is covered by water. A layer of gaseous atmosphere envelopes the earth's surface.

The Earth's Atmosphere

The earth's surface is covered by a layer of atmosphere consisting of a mixture of gases and other solid and liquid particles. The gaseous materials extend to several hundred kilometers in altitude, though there is no well defined boundary for the upper limit of the atmosphere. The first 80 km of the atmosphere contains more than 99% of the total mass of the earth's atmosphere.

Vertical Structure of the Atmosphere

The vertical profile of the atmosphere is divided into four layers: troposphere, stratosphere, mesosphere and thermosphere. The tops of these layers are known as the tropopause, stratopause, mesopause and thermopause, respectively.

- Troposphere: This layer is characterized by a decrease in temperature with respect to height, at a rate of about 6.5ºC per kilometer, up to a height of about 10 km. All the weather activities (water vapour, clouds, precipitation) are confined to this layer. A layer of aerosol particles normally exists near to the earth surface. The aerosol concentration decreases nearly exponentially with height, with a characteristic height of about 2 km.
- Stratosphere: The temperature at the lower 20 km of the stratosphere is approximately constant, after which the temperature increases with height, up to an altitude of about 50 km. Ozone exists mainly at the stratopause. The troposphere and the stratosphere together account for more than 99% of the total mass of the atmosphere.
- Mesosphere: The temperature decreases in this layer from an altitude of about 50 km to 85 km.
- Thermosphere: This layer extends from about 85 km upward to several hundred kilometers. The temperature may range from 500 K to 2000 K. The gases exist mainly in the form of thin plasma, *i.e.* they are ionized due to bombardment by solar ultraviolet radiation and energetic cosmic rays.

The term upper atmosphere usually refers to the region of the atmosphere above the troposphere.

Many remote sensing satellites follow the near polar sun-synchronous orbits at a height around 800 km, which is well above the thermopause.

Atmospheric Constituents

The atmosphere consists of the following components:

- Permanent Gases: They are gases present in nearly constant concentration, with little spatial variation. About 78% by volume of the atmosphere is nitrogen while the life-sustaining oxygen occupies 21%. The remaining one percent consists of the inert gases, carbon dioxide and other gases.
- Gases with Variable Concentration: The concentration of these gases may vary greatly over space and time. They consist of water vapour, ozone, nitrogeneous and sulphurous compounds.
- Solid and liquid particulates: Other than the gases, the atmosphere also contains solid and liquid particles such as aerosols, water droplets and ice crystals. These particles may congregate to form clouds and haze.

10

Platforms and Orbits

INTRODUCTION

High-altitude remote sensing originated in the mid-1800s with aerial photography by balloon and, in at least one instance, the use of cameras attached to the underside of birds. Airplanes became the dominant remote sensing "platform" by the early 20th century. This practice continues to evolve for wartime, intelligence, commercial and government applications. An advantage of airborne remote sensing, is the capability of offering very high spatial resolution images (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, whereas earth observation satellites offer the possibility of continuous monitoring of the earth. The development of satellite remote sensing had greatly improved the ability to cover large areas.

There are two groups of satellites depending on the orbit in which they are placed. A *geostationary* orbit is established when a satellite is placed at a very high altitude, roughly 36,000 km above the earth's equator, and caused to orbit with the earth's rotation (called a *prograde* orbit). The altitude may vary slightly from one geostationary satellite to the next, depending upon the mass of the satellite, but, for the most part, this is a fundamental physical constraint. The rules of geometry-that is, the sight line from the satellite's position above the equator to the farthest edge of the earth's sphere-dictate that geostationary satellites can only "see" limited area at any one time. The laws of physics and capabilities of engineering limit their spatial resolution to a range of about 1 to 10 square kilometres. The total image or scene size, known as the *field of view*, is often thousands of kilometres. Thus, unless a geostationary satellite spins or turns its optics, its view is necessarily fixed. This allows for continuous monitoring, and often a very large, synoptic view of much of one entire hemisphere. The coarse (km range) resolution versus the wide, continuous field of view constitute the main tradeoffs to consider for this orbit type. These characteristics make geostationary satellites best for collecting weather and climate data (such as cloud cover and surface temperature) and relaying communications data, although AVHRR data are used for global-and regional-scale land cover analyses.

The other group of satellites, by far the largest group of earth-orbiting satellites is with the *sun-synchronous* or *polar-orbiting*. These are lunched below the altitude of geostationary satellites closer to the earth's surface, at orbits ranging from 700km to 1000km. These satellites usually orbit at a steep inclination relative to the equator, in the direction opposite the earth's rotation, known as a *retrograde* orbit. When the satellite's orbit and earth's rotation are combined, they result in an s-shaped path relative to a map of the earth's surface. Given enough time, the orbits and rotations of the earth bring the satellite over the same location, leading to the term *exact repeat* satellites. The number of orbits between each return to the same longitude and latitude is called the *repeat cycle*. These satellites usually orbit the earth in roughly 100 to 120 minutes, circling several times per day, returning a satellite to the same position over the earth's surface only after 2 weeks or more. The speed of motion limits the time that a satellite spends over a location, and the amount of time a scanner can "look" at any single ground cell (called the *dwell time*). Most exact repeat satellites that use passive sensors are also in sun synchronous orbits, meaning that they cross the same latitude at the same daylight time with each orbit, but with their location shifted to a different longitude. Their lower altitude allows these satellites to obtain images with spatial resolution ranging from 1-200 meters per side of a pixel, and an image width ranging from tens to thousands of kilometres per scene.

As the satellite passes over the earth's surface, its motion can be described in terms of a ground track that it follows at a certain altitude. Most satellites are *nadir looking*, meaning that their sensing equipment is aimed straight down towards the centre of the earth. However, normal measurements generally include areas substantially on either side of this ground track, and that total area is called the *swath width*. Because most low-orbiting satellites follow a polar orbit, their ground tracks, and thus their swaths, are spread furthest apart at the equator, and are compressed at the poles. As a consequence, there is an overlap, called *side lap*, of neighboring swaths. Logically, this side lap is smallest at the equator and increased at the poles.

A basic understanding of ground tracks, swath widths, and side lap is helpful in designing a remote sensing experiment for a few reasons. First, weather and other temporary conditions may prevent good data acquisition for the first and/or "best" pass of a satellite directly over a given location. In this case, side lap may allow for multiple acquisitions of the same location on the surface with only a short delay, provided that the target location can be seen during two or more subsequent passes of the same satellite. This situation is more likely to occur if the sensing target is at high or low latitude than at the equator. Also, side lap may allow experimenters to obtain data for a single location in much more rapid time series than would be possible if the experimenters were to wait for the satellite to exactly repeat its path-a matter of hours, rather than weeks. Finally, side lap may allow the same location to be viewed from slightly different angles at slightly different times of the same day with neighboring orbital paths. This may provide additional information from shading caused by the sun's angle, and other factors that change with relatively small differences in position and time.

SENSORS AND PLATFORMS

A sensor is a device that measures and records electro-magnetic energy. Sensors can be divided into two groups. Passive sensors depend on an external source of energy, usually the sun. The most common passive sensor is the photographic camera. Active sensors have their own source of energy, an example would be a radar gun. These sensors send out a signal and measure the amount reflected back. Active sensors are more controlled because they do not depend upon varying illumination conditions.

Platforms are 'vehicles' on which sensors are mounted to emit and record EM radiation. Two common platforms in remote sensing are aircrafts (both manned and unmanned) and satellites.

Aircrafts

Aircrafts can be operated at altitudes up to 20 km. They are mostly used to acquire aerial photographs. The advantage of using aircrafts is that high spatial resolution images (down to 5×5 cm pixels) can be acquired for a targeted region at a particular time. The disadvantage is that images normally cover small areas in extent and are expensive to acquire. The use of an aircraft as a platform is normally preceded by flight planning, where pre-defined routes for the aircraft movement, aircraft altitude, dimensions of photographs, etc. are defined. More details on aerial survey missions can be found here.

In recent years, the use of Unmanned Aerial Vehicles (UAVs) has become popular. We refer to the UAV section for a more detailed discussion of the options that this technology provides.

Satellites

When one wants to monitor larger areas (even with global coverage) and/or obtain synoptic views of the same target/area, satellites are normally the platform of choice. Many of the commercial remote sensing images are acquired by sensors on board of satellites. A rocket is used to launch one or more satellites into space where the satellites are placed in a predefined orbit to image the earth for some period of time. Various parameters characterize a satellite's orbit. These include: orbital altitude, inclination angle, period, repeat cycle and type. These orbital parameters to a large extent determine the monitoring capabilities of the satellite. We discuss them below.

Orbital altitude refers to the height (vertical distance) of the satellite above the earth's surface. Most commercial satellites orbit at an altitude of 500–1000 km. Orbital altitude influences, to a very large extent, the spatial coverage and resolution of the resulting image. In most cases, the higher the orbital altitude of a satellite, the larger the spatial coverage (the area that it covers on the earth) and the lower the spatial resolution of the resulting image. Satellites operating from a relatively low orbital altitude are mostly preferred for agricultural applications, especially in smallholder systems, as they provide more spatial detail (they have higher spatial resolution) and this improves chances of identifying the typical small fields. Images from DigitalGlobe satellites, which were mostly used in the STARS project, were taken at orbital altitudes of between 450 and 770 km.

Orbital inclination angle is defined as the angle (in degrees) between the orbital plane and the equatorial plane. Together with the field of view of the sensor, the inclination angle determines the earth's latitudinal extent that a satellite can image. The inclination angle of a satellite is set in accordance with what it was launched to monitor. For example, satellites with very low inclination angles (e.g., 300) may be intended to monitor areas around the tropics only, i.e., areas between latitudes 300 south and 300 north. Well-known commercial sensors such as Landsat and DigitalGlobe's ranges of satellites have orbital inclinations of larger than 900. This means that they observe all areas between latitudes 900 south and 900 north.

Orbital period is the time (in minutes) that it takes a satellite to complete one full orbit. This has implications for the life span of the satellite (i.e., determines number of orbits), but it also determines image quality, temporal resolution and spatial resolution of the resulting images. The orbital period of most commercial sensors ranges between 90 and 100 minutes.

Repeat cycle is defined as the number of days between two successive, identical orbits from a set of identical satellites. In other words, it refers to the frequency at which EO sensors acquire images of the same portion/part of the earth's surface. Repeat cycle is typically measured in days. Satellites that have high repeat cycles (i.e., high acquisition frequency) produce, for the same period, more images than those with a low repeat cycle. For agricultural purposes, satellites with high repeat cycles are preferred, as chances of getting more images during the cropping season are improved. There is always a trade-off between repeat cycle and spatial resolution. Satellites with high repeat cycles mostly produce low spatial resolution images, and vice versa. This problem is gradually being solved by new satellite technology that strives to achieve both. Examples are RapidEye and SPOT6/7.

Orbital type Based on the orbital properties above, three common orbital types can be distinguished: polar, sun-synchronous and geostationary. These are described below.

Polar orbits have an inclination angle between 80 and 1000. By virtue of this angle, these satellites observe virtually the whole globe. DigitalGlobe's range of satellites falls within this category, and so do the Landsat satellites and many other commercial satellites. Satellites with this orbital type normally have an orbital altitude between 600 and 1000 km.

Sun-synchronous orbits appear to be in the same position from the perspective of the sun always. Satellites in sun-synchronous orbits have high inclination angle and cross the equator at the same local time in every orbit. Most satellites in this orbit cross the equator at around 10.30 am local time. This is probably an optimal timing for combination of good sun illumination and cloud cover. The satellite has the capability to record night time images (which results in thermal bands) on the return leg of its orbit. Landsat, SPOT, IRS and DigitalGlobe's range of satellites are all in sun-synchronous orbits.

Geostationary: satellites that have an inclination angle of zero (i.e., they are placed above the equator) and orbital altitude of an approximate (and very high) 36,000 km are said to be in geostationary orbit. Satellites in this category have an orbital period of one sidereal day, which is equal to the rotational period of the earth around its own axis. The altitude of these satellites leads to low spatial resolution but high spatial coverage for the resulting images. Most satellites in this orbital type category are used for meteorological and telecommunication purposes. Images from such satellites are not directly useful for smallholder agricultural applications, though obviously weather data is important to farmers.

ORBITS AND SWATHS

The path followed by a satellite is referred to as its orbit. Satellites which view the same portion of the earth's surface at all times have Geostationary orbits.Weather and communication satellites commonly have these types of orbits.

Many satellites are designed to follow a north south orbit which, in conjunction with the earth's rotation (west-east), allows them to cover most of the earth's surface over a period of time. These are Near-polar orbits. Many of these satellites orbits are also Sun-synchronous such that they cover each area of the world at a constant local time of day.

Near polar orbits also means that the satellite travels nort-ward on one side of the earth and the southward on the second half of its orbit. These are called Ascending and Descending passes. As a satellite revolves around the earth, the sensor sees a certain portion of the earth's surface. The area imaged is referred to as the Swath. The surface directly below the satellite is called the Nadir point. Steerable sensors on satellites can view an area (off nadir) before and after the orbits passes over a target.

SATELLITE SENSOR CHARACTERISTICS

The basic functions of most satellite sensors is to collect information about the reflected radiation along a pathway, also known as the field of view (FOV), as the satellite orbits the Earth. The smallest area of ground that is sampled is called the instantaneous field of view (IFOV). The IFOV is also described as the pixel size of the sensor. This sampling or measurement occurs in one or many spectral bands of the EM spectrum. The data collected by each satellite sensor can be described in terms of spatial, spectral and temporal resolution.

SPATIAL RESOLUTION

The spatial resolution (also known as ground resolution) is the ground area imaged for the instantaneous field of view (IFOV) of the sensing device. Spatial resolution may also be described as the ground surface area that forms one pixel in the satellite image. The IFOV or ground resolution of the Landsat Thematic Mapper (TM) sensor, for example, is 30 m. The ground resolution of weather satellite sensors is often larger than a square kilometer. There are satellites that collect data at less than one meter ground resolution but these are classified military satellites or very expensive commercial systems.

TEMPORAL RESOLUTION

Temporal resolution is a measure of the repeat cycle or frequency with which a sensor revisits the same part of the Earth's surface. The frequency will vary from several times per day, for a typical weather satellite, to 8—20 times a year for a moderate ground resolution satellite, such as Landsat TM. The frequency characteristics will be determined by the design of the satellite sensor and its orbit pattern.

PLATFORMS

Aerial photography has been used in agricultural and natural resource management for many years. These photographs can be black and white, colour, or colour infrared. Depending on the camera, lens, and flying height these images can have a variety of scales. Photographs can be used to determine spatial arrangement of fields, irrigation ditches, roads, and other featuresor they can be used to view individual features within a field. Infrared images can detect stress in crops before it is visible with the naked eye.

Healthy canopies reflect strongly in the infrared spectral range, whereas plants that are stressed will reflect a dull colour. These images can tell a farmer that there is a problem but does not tell him what is causing the problem. The stress might be from lack of water, insect damage, improper nutrition or soil problems, such as compaction, salinity or inefficient drainage. The farmer must assess the cause of the stress from other information. If the dull areas disappear on subsequent pictures, the stress could have been lack of water that was eased with irrigation. If the stress continues it could be a sign of insect infestation. The farmer still has to conduct in-field assessment to identify the causes of the problem.

The development of cameras that measure reflectance in a wider range of wavelengths may lead to better quantify plant stress. The use of these multi-spectral cameras are increasing and will become an important tool in precision agriculture. Satellite remote sensing is becoming more readily available for use in precision agriculture. The Landsat and the NOAA polar-orbiting satellites carry instruments that can be used to determine crop types and conditions, and to measure crop acreage.

The Advanced Very High Resolution Radiometer (AVHRR) carried onboard NOAA polar orbiting satellites measure reflectance from the earth's surface in the visible, near infrared, and thermal infrared portions of the electromagnetic spectrum. This spectral sensitivity makes it suitable for measuring vegetative condition and because the satellite passes overhead twice a day, it can be used to detect rapidly changing conditions.

Unfortunately, its use as a precision agriculture tool is limited because the spatial resolution of the sensor is nominally 1.1km. A possible application of this scanner would be to use the thermal infrared sensor to estimate daily maximum and minimum temperatures. These temperature estimates could then be used to determine degree-days that will drive pest development models. Degree-day models are an essential part of IPM programmes and the enhanced spatial coverage provided by satellites would allow for assessment of spatial variability in predicted events that is not possible with data from sparsely spaced weather stations currently used for these models.

Remotely sensed data can also be used to determine irrigation scheduling and adequacy of irrigation systems for uniformly wetting an entire field. The sensors aboard the Landsat satellite measures reflected radiation in seven spectral bands from the visible through the thermal infrared. The sensors high spatial resolution (approximately 30m) makes it useful in precision agriculture.

The spectral response and higher spatial resolution make it suitable for assessing vegetative condition for individual fields but the overpass frequency is only once every 16 days. The less frequent overpass makes it difficult to use these data for assessing rapidly changing events such as insect outbreaks or water stress. New satellites with enhanced capabilities are planned and remotely sensed data will become more widely used in management support systems.

COMMON SATELLITES

GOES: 5 spectral bands 1-41 km spatial resolution Geostationary NOAA AVHRR: 5 spectral bands 1.1 km spatial resolution 1 day repeat cycle Landsat TM: 7 spectral bands 30m spatial resolution 16 day repeat cycle MODIS: Multi-spectral bands

250-1000m spatial resolution (band dependent) 1day repeat cycle
IKONOS: 4 spectral Bands 4m spatial resolution 5 day repeat cycle

Remote sensing makes use of visible, near infrared and short-wave infrared sensors to form images of the earth's surface by detecting the solar radiation reflected from targets on the ground. Different materials reflect and absorb differently at different wavelengths. Thus, the targets can be differentiated by their spectral reflectance signatures in the remotely sensed images.

SPECTRAL REFLECTANCE SIGNATURE

When solar radiation hits a target surface, it may be transmitted, absorbed or reflected. Different materials reflect and absorb differently at different wavelengths. The reflectance spectrum of a material is a plot of the fraction of radiation reflected as a function of the incident wavelength and serves as a unique signature for the material. In principle, a material can be identified from its spectral reflectance signature if the sensing system has sufficient spectral resolution to distinguish its spectrum from those of other materials. This premise provides the basis for multispectral remote sensing. The following graph shows the typical reflectance spectra of water, bare soil and two types of vegetation.

The reflectance of clear water is generally low. However, the reflectance is maximum at the blue end of the spectrum and decreases as wavelength increases. Hence, water appears darkbluish to the visible eye. Turbid water has some sediment suspension that increases the reflectance in the red end of the spectrum and would be brownish in appearance. The reflectance of bare soil generally depends on its composition.

In the example shown, the reflectance increases mono-tonically with increasing wavelength. Hence, it should appear yellowish-red to the eye. Vegetation has a unique spectral signature that enables it to be distinguished readily from other types of land cover in an optical/near-infrared image. The reflectance is low in both the blue and red regions of the spectrum, due to absorption by chlorophyll for photosynthesis. It has a peak at the green region. In the near infrared (NIR) region, the reflectance is much higher than that in the visible band due to the cellular structure in the leaves.

Hence, vegetation can be identified by the high NIR but generally low visible reflectance. This property has been used in early reconnaissance missions during war times for "camouflage detection". The shape of the reflectance spectrum can be used for identification of vegetation type. For example, the reflectance spectra of dry grass and green grass in the previous figures can be distinguished although they exhibit the generally characteristics of high NIR but low visible reflectance. Dry grass has higher reflectance in the visible region but lower reflectance in the NIR region. For the same vegetation type, the reflectance spectrum also depends on other factors such as the leaf moisture content and health of the plants. These properties enable vegetation condition to be monitored using remotely sensed images.

11

Sensors and Scanning Systems

CHARACTERISTICS OF REMOTE SENSORS

A sensor is characterized by its spectral properties (number and placement of bands), its orbital altitude and path, its swath width, and its *spatial resolution*.

Several different vector data models exist, however only two are commonly used in GIS data storage. The topologic data structure is often referred to as an *intelligent data structure* because spatial relationships between geographic features are easily derived when using them. Primarily for this reason the topologic model is the dominant vector data structure currently used in GIS technology. Many of the complex data analysis functions cannot effectively be undertaken without a topologic vector data structure.

Spatial resolution is measured in terms of the size of one pixel projected on the ground. Spatial resolution is directly tied to the size of the features that can be resolved (or "seen") on the ground. The higher the resolution, the less likely that there will be "mixed pixels" in which radiances effectively represent an average of land cover types in the ground area represented by that pixel (*e.g.*, half lake and half forest). Commercial high resolution sensors have a spatial resolution in the 0.6-10 meter range, medium resolution sensors fall in the 10-50 meter range, and low resolution sensors have greater than 50 meter resolution.

Until the advent of the commercial satellites IKONOS and QuickBird, with resolutions of one square meter or finer, high resolution imagery was the exclusive province of intelligence-gathering agencies. Most social science applications do not command the financial resources required to obtain such high resolution data, nor are images of this resolution generally required, except perhaps in the area of international relations, law and policy. Most social science research tends to utilize polar-orbiting satellites with medium spatial resolution, such as the Landsat, SPOT, TERRA and, more recently, AQUA satellites. They provide good spectral and ground resolution, with multiple visible, infrared, and panchromatic bands and pixel width ranging from 5 to 30 meters.

Data Processing, Interpretation and Analysis

Much of the technical work of remote sensing involves pre-processing and applying radiometric and geometric corrections to imagery to compensate for errors due to factors such as atmospheric interference of incoming radiation and sensor and data stream irregularities. Once such corrections are applied, imagery must be *georeferenced*. Georeferencing is the process of taking the image in its raw format (rows and columns of data) and linking it to the land that it covers. Images are georeferenced by linking spatially distributed control points in the satellite image to points on base maps or points referenced in the field through global positioning systems. The raster data in the image is thereby registered to a Cartesian coordinate system, and can be combined with other georeferenced data sets in a geographic information system.

The processed data can now either be visually interpreted or classified using manual or automated processes. The main elements of visual image interpretation involve gradients of tone or colour, resolution, size and shape, texture and pattern, site and association, and height and shadows. Given their knowledge of the characteristic spectral signatures of different land cover types, scientists my inspect black and white images of each band separately in order to identify features and patterns. For many purposes, data that is collected from the earth's surface, which represents a continuous variation, needs to be categorized (de Sherbinin *et al.* 2002). Image classification is the process of creating discrete classes or categories of land cover, utilizing information from some or all of the bands to group together pixels with similar spectral signatures. Supervised classification entails providing the software with sample pixels that represent specific features, such as boreal forest, and then having the computer classify every pixel with a similar *spectral signature* as boreal forest. Analysts may also use images from different seasons in order to discriminate vegetation cover types that have different phenologies, such as deciduous and evergreen forests. In unsupervised classification, the analyst specifies the desired number of classes, and the computer automatically sorts the pixels according to their spectral signatures. The analyst then labels the resulting groups based on some local knowledge of the land cover patterns.

Once classified, it is necessary to verify or validate that the output product accurately represents the actual composition, content, structure or land surface characteristics being mapped. Validation requires either field visits, ground-truthing or comparing the classified image with existing maps or images of sufficient detail. Statistics can be derived for the classified imagery indicating the general and specific (class-wise) agreement between the pixels or classes used, letting the user know which were classified correctly and which ones were not.

Validation results are also sometimes presented as a percentage value associated with the map that communicates how accurate the map is on a per pixel basis. Since the highest confidence rankings reported by satellite land cover data sets are between 85% and 90% (for the easiest types of land cover to classify), for an image with a per pixel accuracy of 85% the probability that one pixel out of four is incorrectly classified is close to 0.50. The output of remote sensing data analysis can be presented in a variety of ways including a printout of the enhanced image itself a spatial database, summary statistics and/or graphs (Jensen 1996). The output data can be integrated with a geographic information system (GIS) database for further analysis.

SENSOR SYSTEMS

Platforms

Platforms refer to the structures or vehicles on which remote sensing instruments are mounted. The platform on which a particular sensor is housed determines a number of attributes, which may dictate the use of particular sensors. These attributes include: distance the sensor is from the object of interest, periodicity of image acquisition, timing of image acquisition, and location and extent of coverage. There are three broad categories of remote sensing platforms: ground based, airborne, and satellite.

Ground based -- A wide variety of ground based platforms are used in remote sensing. Some of the more common ones are hand held devices, tripods, towers and cranes. Instruments that are ground-based are often used to measure the quantity and quality of light coming from the sun or for close range characterization of objects. For example, to study properties of a single plant or a small patch of grass, it would make sense to use a ground based instrument.

Laboratory instruments are used almost exclusively for research, sensor calibration, and quality control. Much of what is learned from laboratory work is used to understand how remote sensing can be better utilized to identify different materials. This contributes to the development of new sensors that improve on existing technologies.

Field instruments are also largely used for research purposes. This type of remote sensing instrument is often hand-held or mounted on a tripod or other similar support. 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.

The BOREAS (BOReal Ecosystem-Atmosphere Study) field experiment was conducted to gain knowledge about relationships between the boreal forest and Earth's atmosphere.

Airborne -- Airborne platforms were the sole non-ground-based platforms for early remote sensing work. The first aerial images were acquired with a camera carried aloft by a balloon in 1859. Balloons are rarely used today because they are not very stable and the course of flight is not always predictable, although small balloons carrying expendable probes are still used for some meteorological research.

At present, airplanes are the most common airborne platform. Nearly the whole spectrum of civilian and military aircraft are used for remote sensing applications. When altitude and stability requirements for a sensor are not too demanding, simple, low-cost aircraft can be used as platforms. However, as requirements for greater instrument stability or higher altitudes become necessary, more sophisticated aircraft must be used.

In this section, aircraft are divided into three categories (low, mid, and high) based on their altitude restrictions. In general, the higher an aircraft can fly, the more stable a platform it is, but correspondingly more costly to operate and maintain.

Low altitude aircraft typically fly below altitudes where supplemental oxygen or pressurization are needed (12,500 feet above sea level). They are good for acquiring high spatial resolution data limited to a relatively small area. Included in this class are the common fixed-wing, propeller driven planes used by private pilots, such as the Cessna 172 or 182, and Piper Cherokee. This class of aircraft is inexpensive to fly and can be found throughout the world. Some of these airplanes are specially outfitted for mounting remote sensing instruments in the underside of the plane, however, many times instruments are simply hung out the door using simple mounts.

Helicopters are usually used for low altitude applications where the ability to hover is required. Helicopters are quite expensive to operate and they are typically used only when needed. Ultralight aircraft are a class of aircraft that is gaining popularity.

The Federal Aviation Authority (FAA) defines an ultralight as a single seat powered flying machine that weighs less than 254 pounds, has a top speed of 55 knots (63 mph), stalls at 24 knots (28 mph) or less and carries no more than 5 gal. of fuel. These small, often portable, aircraft are inexpensive and are able to take off and land where larger aircraft cannot. They are limited to flying at lower elevations and at slow speeds. If the demands of the remote sensing requirement are not too strict, ultralight aircraft may be a reasonable alternative to larger aircraft.

Midaltitude aircraft have an altitude limit under 30,000 feet above sea level. This includes a number of turbo-prop aircraft. Often at higher altitudes, there is less turbulence so stability is better. This class of airplane is used when stability is more important and when it is necessary or desired to acquire imagery from a greater distance than available from low altitude aircraft.

High altitude aircraft can fly at altitudes greater than 30,000 feet above sea level. This class of airplane is usually powered by jet engines and is used for specialized tasks, such as atmospheric studies, research to simulate satellite platforms, and other applications where a high altitude platform is required. High altitude aircraft are good for acquiring large areal coverage with typically lower spatial resolutions.

Another class of aircraft that has been in use for many years is remote control aircraft, or drones. Remotely controlled aircraft are often used for conditions when it may be too hazardous to fly. They have been used extensively by the military.

Satellite -- The most stable platform aloft is a satellite, which is spaceborne. The first remote sensing satellite was launched in 1960 for meteorology purposes. Now, over a hundred remote sensing satellites have been launched and more are being launched every year. The Space Shuttle is a unique spacecraft that functions as a remote sensing satellite and can be reused for a number of missions.

Satellites can be classified by their orbital geometry and timing. Three orbits commonly used for remote sensing satellites are geostationary, equatorial and Sun synchronous. A geostationary satellite has a period of rotation equal to that of Earth (24 hours) so the satellite always stays over the same location on Earth. Communications and weather satellites often use geostationary orbits with many of them located over the equator. In an equatorial orbit, a satellite circles Earth at a low inclination (the angle between the orbital plane and the equatorial plane). The Space Shuttle uses an equatorial orbit with an inclination of 57 degrees.

Sun synchronous satellites have orbits with high inclination angles, passing nearly over the poles. Orbits are timed so that the satellite always passes over the equator at the same local sun time. In this way the satellites maintain the same relative position with the sun for all of its orbits. Many remote sensing satellites are Sun synchronous which ensures repeatable sun illumination conditions during specific seasons. Because a Sun synchronous orbit does not pass directly over the poles, it is not always possible to acquire data for the extreme polar regions. The frequency at which a satellite sensor can acquire data of the entire Earth depends on sensor and orbital characteristics. For most remote sensing satellites the total coverage frequency ranges from twice a day to once every 16 days.

Another orbital characteristic is altitude. The Space Shuttle has a low orbital altitude of 300 km whereas other common remote sensing satellites typically maintain higher orbits ranging from 600 to 1000 km.

Most remote sensing satellites have been designed to transmit data to ground receiving stations located throughout the world. To receive data directly from a satellite, the receiving station must have a line of sight to the satellite. If there are not sufficient designated receiving stations around the world, any given satellite may not readily get a direct view to a station, leading to potential problems of data discontinuity. To work around this problem, data can be temporarily stored onboard the satellite and then later downloaded upon acquiring contact with the receiving station. Another alternative is to relay data through TDRSS (Tracking and Data Relay Satellite System), a network of geosynchronous (geostationary) communications satellites deployed to relay data from satellites to ground stations.

The payload for remote sensing satellites can include photographic systems, electro-optical sensors, microwave or lidar systems. For applications benefiting from simultaneous coverage by different sensors, more than one sensing system can be mounted on a single satellite. In addition to sensor systems, there are often devices for recording, preprocessing and transmitting the data.

FUNDAMENTAL SENSOR TYPES

There are several broad categories of basic sensor system types such as passive vs. active, and imaging vs. nonimaging. Passive vs. active refers to the illumination source of the system; imaging vs. nonimaging refers to the form of the data. A variety of different sensors fit in these categories, which are not mutually exclusive.

Passive vs. active sensors -- Passive sensors measure light reflected or emitted naturally from surfaces and objects. Such instruments merely observe, and depend primarily on solar energy as the ultimate radiation source illuminating surfaces and objects. Active sensors (such as radar and lidar systems) first emit energy (supplied by their own energy source) and then measure the return of that energy after it has interacted with a surface. Use of data collected by passive sensors often requires accurate measurements of solar radiation reaching the surface at the time the observations were made. This information allows for the correction of "atmospheric effects" and results in data or images that are more representative of actual surface characteristics.

Imaging vs. nonimaging sensors -- Remote sensing data are the recorded representation of radiation reflected or emitted from an area or object. When measuring the reflected or emitted energy, either imaging or nonimaging sensors can be used. Data from imaging sensors can be processed to produce an image of an area, within which smaller parts of the sensor's whole view are resolved visually (see discussion of pixels below). Nonimaging sensors usually are hand held devices that register only a single response value, with no finer resolution than the whole area viewed by the sensor, and therefore no image can be made from the data. These single values can be referred to as a type of "point" data, however some small area is typically involved depending on the sensor's spatial resolution.

Image and nonimage data each have particular uses. Nonimage data give information for one specific (usually small) area or surface cover type, and can be used to characterize the reflectance of various materials occurring in a larger scene and to learn more about the interactions of electromagnetic energy and objects. Image data provide an opportunity to look at spatial relationships, object shapes, and to estimate physical sizes based on the data's spatial resolution and sampling. Image data are desirable when spatial information (such as mapped output) is needed. This text refers primarily to imaging sensors and data.

Images produced from remote sensing data can be either analog (such as a photograph) or digital (a multidimensional array or grid of numbers). Digital data can be analyzed by studying the values using calculations performed on a computer, or processed to produce an image for visual interpretation. Image interpretation is used to decipher information in a scene. In the past, image interpretation was done largely using subjective visual techniques, but with the development and ongoing advancement of computer technology, numeric or digital processing has become a powerful and common interpretation tool.

In many cases, image interpretation involves the combination of both visual and digital techniques. These techniques utilize a number of image features including tone and color, texture, shape, size, patterns, and associations of objects. The human eye and brain are generally thought to more easily process the spatial characteristics of an image, such as shape, patterns and how objects are associated with one another. Computers usually are better suited for rapid analysis of the spectral elements of an image such as tone and color. Sophisticated computer software that can perform like the human eye and brain may be more commonly available in the future.

Passive Sensors

Passive sensors are the most common sensor type for vegetation related remote sensing. This is not only because passive sensor systems are generally simpler in design (built only to receive energy) but also because portions of the solar spectrum provide very useful information for monitoring plant and canopy properties.

A major limitation of passive systems is that in most cases they require sunlight in order for valid and useful data to be acquired. Consequently, deployment of or data acquisition by passive sensors is very dependent on lighting (time of day, time of year, latitude) and weather conditions, since cloud cover can interfere with the path of solar radiation from the sun to the surface and then to the sensor.

The signals detected by passive sensors can be greatly altered due to atmospheric effects, especially in the shorter wavelengths of the solar spectrum that are strongly scattered by the atmosphere. These effects can be minimized (but not eliminated) by collecting data only under very clear and dry atmospheric conditions.

Sophisticated atmospheric correction routines now exist to remove atmospheric effects from data acquired by passive sensors.

Photographic -- The most common sensor system is the photographic camera -- a simple passive sensor. Many of the historic developments in remote sensing were directly related to the development of photographic systems. Camera systems are similar in design to the human eye. Both have a lens at one end of an enclosed chamber and a light-sensitive material (film for a camera and the retina for an eye) at the other. In both systems, an iris is used to control the amount of light that can strike the film/retina. In a camera, a shutter is placed between the lens and film to control how long the light can strike the film. Filters can be attached in front of a lens to restrict the wavelength of light permitted to strike the film.

There are three basic elements of photographic systems -- optics, film, and filters. Optics refer to lenses and the geometry of light retrieval in a camera. The lenses in a camera are responsible for focusing and zooming on an object. Before light reflected from an object strikes the film, it must pass through one or more lenses. As light passes through a lens, it is bent to focus the imaged object on the film. To minimize distortions associated with the use of single lenses, most camera lenses are actually composed of multiple lenses that work in concert to form an image onto the film.

The amount of image detail that can be recorded on film is directly related to the distance between the lens and the film, referred to as the focal length. As the focal length increases, the detail that can be seen on the film increases. Increasing the focal length is commonly called zooming in on an object.

Film in a camera is used to record the image that passes through the lens. Photographic film is composed of a durable base, which is coated with a light-sensitive layer know as the emulsion. During the short time that a shutter is open, light strikes the film and leaves a latent image on the emulsion. This image can be made visible by the process of developing and printing.Emulsions are made of materials sensitive to particular regions of the electromagnetic spectrum. For example, some film is only sensitive to visible light, whereas other film is sensitive to near-infrared light. In color film, the emulsion is composed of three layers, with each being sensitive to different wavelengths of light, normally blue, green and red light. With black and white film, the emulsion is sensitive to a broad spectrum of light. Film emulsions are generally limited to recording wavelengths between 0.4 to 0.9 micrometers.

Film speed is another quality of emulsions that is important for aerial photography. Film speed refers to the quantity of light than is needed to expose the emulsion. Fast film requires less light than slow film to record the same image. If the camera platform is moving, one would want to use a high speed film to reduce the blurring effects of the moving camera. Unfortunately, there is a tradeoff between film speed and image quality -- the faster the film speed, the grainier the image. Because of this tradeoff, it is necessary to carefully choose a film speed that will meet the requirements of the end user. Some sophisticated camera mounts have an image motion compensator that reduces the blurring effect of the moving platform, which potentially allows the use of slower film.

In many remote sensing applications, it is important to restrict the light entering the camera by the use of filters. Color filters work by absorbing a range of wavelengths while allowing other wavelengths to pass through. Another filter type, known as neutral color filters, do not alter the spectral composition of light, but instead reduce the amount of light of all wavelengths that pass through.

Perhaps the most common color filter is an antihaze filter. These are clear or yellow filters, which absorb out the shorter ultraviolet and blue wavelengths that are substantially scattered by particulates in the atmosphere. Another filter used for monitoring vegetation is an infrared filter, which absorbs visible light and only allows infrared light to pass through.

Aerial photography is one of the oldest forms of remote sensing and it is still used extensively today. It is usually the choice if great spatial detail is needed. For example, photography can be used to identify individual tree species (based on the shape of individual trees) and measure tree heights using special photographic techniques. Because of the detail that can be discerned on a photograph, aerial photography is used extensively for mapping vegetation classes.

Aerial photography is also used as a reconnaissance tool to provide overview information for a particular area. For instance, if there has been an outbreak of a disease that is killing a certain tree or agricultural species, aerial photography using infrared film (to locate trees that are being stressed) can monitor areas for signs and extent of the disease.

Electro-optic radiometers -- A radiometer is an instrument designed to measure the intensity of electromagnetic radiation in a set of wavebands ranging from the ultraviolet to microwave wavelengths. Microwave radiometers are discussed in Section 3.3.3. The radiometers discussed here are called electro-optic sensors because they measure electromagnetic energy using optical techniques and electronic detectors. Though they are only capable of recording a single data value for their view area, if they are mounted in a scanner device images can be produced.

Radiometers are similar in design to a camera in that they have an opening for the light to enter, lenses and mirrors for the light to pass through, but instead of film, they have an electronic detector to record the intensity of electromagnetic energy. As energy hits the detector, a signal proportional to the incoming irradiance is processed to either a digital or analog output that can be recorded.

Detectors for radiometers have been devised to measure wavelengths from 0.4 to 14 micrometers. Although some radiometers can detect this entire range of wavelengths, most only measure selected wavebands in this range. Radiometers that measure more than one waveband are called multispectral radiometers. For this type of radiometer, the light must be separated into discrete wavebands so that multiple waveband or multichannel readings can be taken. This separation can be done using filters, prisms or other sophisticated techniques.

Nonimaging radiometers are commonly used as research tools to better understand how light interacts with objects, for spectral characterization of a variety of surfaces, and for atmospheric measurements. Another common use is to measure the quantity and quality of solar energy. These measurements can in turn be used to correct other imaging and nonimaging measurements for atmospheric effects.

Passive microwave systems -- Passive microwave systems are based on a type of radiometer that detects wavelengths in the microwave region of the spectrum. Because of the nature of microwave radiation, optical systems cannot be used for the detection of this range of wavelengths. As with optical systems though, both nonimaging and imaging systems are available. The components of a microwave radiometer are an antenna, receiver, and recording device. Microwave energy emitted from Earth's surface is collected by an antenna, converted by a receiver into a signal, and recorded.

The features of electromagnetic energy measured by microwave radiometers are polarity, wavelength, and intensity. These properties provide useful information about the structure and composition of an object. Most of the applications of passive microwave radiometers have been in the fields of atmospheric and oceanographic research. It has also proven to be an effective tool for the measurement of soil moisture, an important parameter in studying vegetation.

Visible, infrared, and thermal imaging systems -- By combining a number of detectors or radiometers into detector arrays, it is possible to create a sensor that can acquire a 2D image of an area. There are three basic designs for imaging sensors: frame, pushbroom, and mechanical scanner.

The first two designs are similar. The frame sensor is a 2D array of detectors that acquires an entire image in one exposure similar to the way a camera captures an image on film. A pushbroom sensor is a 1D array that obtains an image one line at a time. Each new data line is added as the platform moves forward, building up an image over time. In a mechanical scanner system the sensor acquires only one or several pixels in any given instant, but since the scanner physically sweeps or rotates the sensor (a radiometer) or a mirror back and forth, an image is produced.

This category of sensor (passive visible, infrared and thermal imaging systems) contains numerous instruments that have been deployed on a wide variety of platforms and used for many applications. Most modern imaging systems are multispectral (acquiring data for more than one limited spectral area). The recording of each discrete spectral sampling is referred to as an image band or channel. Using image processing techniques, multiple (usually three) bands selected from a multispectral image database can be combined to make a single color composite image.

ACTIVE SENSORS

Active systems supply their own illumination energy which can be controlled. Some advantages active systems have over passive sensors are they do not require solar illumination of surfaces or perfect weather conditions to collect useful data. Consequently they can be deployed at night or in conditions of haze, clouds, or light rain (depending on the wavelength of the system).

Radar (active microwave) -- Radar (radio detection and ranging) systems use microwaves (wavelengths ranging from 1 millimeter to 1 meter). Microwave pulses are transmitted at a target or surface, and the timing and intensity of the return signal is recorded.

Transmission characteristics of radar depend on the wavelength and polarization of the energy pulse. Common wavelength bands used in pulse transmission are K-band (11-16.7 mm), X-band (24-37.5 mm), and L-band (150-300 mm). The use of letter codes to designate the wavelength range for various radar systems originated when radar was being developed during World War II. The random letter designations were assigned arbitrarily to ensure military security, however their use has persisted. Distinct from wavelength is the polarization of the transmitted energy. Pulses can be transmitted or received in either an H (horizontal) or V (vertical) plane of polarization.

Factors determining the strength of a radar return signal are complex and varied, however the most important are geometric and electrical properties of the surface or object that reflects the signal.

Information about the structure and composition of objects and surfaces can be detected with radar. Radar has been used in a number of fields, including geology, snow and ice studies, oceanography, agriculture, and vegetation studies. Radar has been especially useful in areas with nearly constant cloud cover.

Lidar (active optical) -- Lidar (light detecting and ranging) systems use laser light as an illumination source. A short pulse of light is emitted from a laser and a detector receives the light energy (photons) after it has been reflected, or absorbed and remitted, by an object or surface. Lidar systems emit pulses at specific, narrow wavelengths that depend on the type of laser transmitter used. The possible wavelengths range from about 0.3 to 1.5 micrometers, which covers the ultraviolet through near-infrared spectral range. The simplest lidar systems measure the round trip travel time of a laser pulse, which is directly related to the distance between the sensor and the target. Basic distance measuring lidars are often referred to as rangefinders or as laser altimeters if deployed on an aircraft or spacecraft. These systems typically measure elevation, slope, and roughness of land, ice, or water surfaces.

More advanced lidars measure the received intensity of the backscattered light as a function of travel time. The intensity of the signal provides information about the material that reflected the photons. Such backscatter lidar systems are often used for atmospheric monitoring applications concerned with the detection and characterization of various gases, aerosols and particulates. Lidar methods have recently been adapted to measure tree heights and the vertical distribution of canopy layers with great accuracy and precision. Lidar instruments have flown on the Space Shuttle, and Vegetation Canopy Lidar (VCL) and Ice, Cloud, and land Elevation Satellite (ICESat) lidar missions are planned for the near future.

Lidar systems can also make fluorescence measurements. Fluorescence refers to the process where a material absorbs radiant energy at one wavelength and then emits it at a different wavelength without first converting the absorbed energy into thermal energy. The wavelengths at which absorption and emission occur are specific to particular molecules. Fluorescence data can identify and quantify the amount of plankton and pollutants in the marine environment. Leaf fluorescence can also help to identify plant species.

REMOTE SENSING SOFTWARE

Remote Sensing data is processed and analysed with computer software, known as a remote sensing application. A large number of proprietary and open source applications exist to process remote sensing data. According to an NOAA Sponsored Research by Global Marketing Insights, Inc. the most used applications among Asian academic groups involved in remote sensing are as follows: ERDAS 36% (ERDAS IMAGINE 25% & ERMapper 11%); ESRI 30%; ITT Visual Information Solutions ENVI 17%; MapInfo 17%. Among Western Academic respondents as follows: ESRI 39%, ERDAS IMAGINE 27%, MapInfo 9%, AutoDesk 7%, ITT Visual Information Solutions ENVI 17%. Other important Remote Sensing Software packages include: TNTmips from MicroImages, PCI Geomatica made by PCI Geomatics, the leading remote sensing software package in Canada, IDRISI from Clark Labs, Image Analyst from Intergraph, and the original object based image analysis software eCognition from Definiens. Dragon/ips is one of the oldest remote sensing packages still available, and is in some cases free. Open source remote sensing software includes GRASS GIS, QGIS, OSSIM, Opticks (software) and Orfeo toolbox.

SCANNER PLATFORM SYSTEMS

Airplanes have served as remote sensing platforms starting with Wilber Wright carrying the first camera into the air. Aircraft have several useful advantages as platforms for remote sensing systems. Aircraft can fly at relatively low altitudes thus allowing for sub-meter sensor spatial resolution. Aircraft can easily change their schedule to avoid weather problems such as clouds, which may block a passive sensor's view of the ground. Last minute timing changes can be made to adjust for illumination from the sun, the location of the area to be visited and additional revisits to that location. Sensor maintenance, repair and configuration changes are easily made to aircraft platforms. Aircraft flight paths know no boundaries except political boundaries. Getting permission to intrude into foreign airspace can be a lengthy and frustrating process. The low altitude flown by aircraft narrows the field of view to the sensor requiring many passes 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 of returning the aircraft to the airport before transferring the raw image data to the data provider's facility for preprocessing.

Satellite Systems

Satellite platforms flown from space provide a very wide field of view for the sensor and regular systematic repetitive revisits. Resolution is limited due to the satellite's fixed altitude and orbital path flown. Satellites know no political boundaries allowing them to cover any corner of the globe unheeded by foreign government interference. Expensive ground support facilities are required to operate satellites. The satellites systems are capital intensive costing hundreds of millions of dollars and have relatively short operating life spans of usually five years or less.

Major Satellite Programs

Some major satellite programs delivering images used in agriculture today include the following:

Landsat 5 uses a thematic sensor ("TM") which operates in 7 bands with a resolution of 30 meters except thermal infrared with has a resolution of 120 meters. Space Imaging EOSAT of Thornton, Colorado is the exclusive distributor of Landsat images.

Spot 1,2,3, and 4 use high-resolution visible ("HRV") sensors that operate in 4 bands with a resolution of 10 m panchromatic and 20 m multispectral. Spot images are distributed by Spot Image headquartered in Toulouse, France.

IRS-1C uses three sensors: the LISS-III, with 23 meter resolution in four spectral bands, a panchromatic sensor, with 5.8 m resolution, and a Wide Field Sensor ("WiFS"), with 188 m resolution. IRS images are distributed by Space Imaging EOSAT of Thornton, Colorado under an exclusive license from ANTRIX Corp. Ltd. of India, the commercial marketing company of the Indian Space Research Organization. Investing in commercial satellites can be a risky business. TRW's Lewis satellite with hyperspectral sensors was lost shortly after launch in August of 1997. EarthWatch also lost its EarlyBird satellite four days after launch in December of 1997.

Terrestrial Systems

Terrestrial remote sensing systems are ground-based sensor systems. Some research has been done using remote sensors attached to long hydraulic booms hoisted above the crop canopy from the ground. Images collected from such a close distance have resolutions that are much greater than images from aircraft or satellites. Other ground-based systems use vehicle-mounted sensors that control variable rate applicators in real time. For example, remote sensors that can distinguish weeds from the crop are mounted on sprayers that change the application rate of herbicides applied on the go. A form of remote sensing technology called *machine vision* is used to sense weeds in the crop and control the sprayer.

The World Agricultural Outlook Board ("WAOB") is one of the U.S. Federal Government's largest users of remote sensing (USDA WAOB, 1998). The WAOB coordinates all remote sensing activities for the United States Department of Agriculture ("USDA") agencies. USDA agencies use remote sensing to assess crop conditions; monitor, manage and administer natural resources; and conduct remote sensing research.

The USDA provides statistical information about agriculture and rural communities through its agency the National Agricultural Statistics Service ("NASS"). The NASS conducts surveys and prepares estimates of U.S. agricultural crop production, supply inventories and agricultural production revenue and costs. Federal, State and Local governments use this information to help form public policy and legislation controlling agricultural commodity production, storage, marketing, and distribution.

The NASS has had limited success using remote sensing to assist in estimating crop yields. NASS statisticians use Landsat satellites to identify patterns in crops in order to assess crop development and progress. Problems cited by the NASS for limiting its remote sensing to only one or two states include temporal availability that doesn't allow for an adequate number of revisits and cloud cover similar that has reduced or blocked the number of useful images. Every two weeks, NASS statisticians use remote sensing information from the National Oceanic and Atmospheric Administration ("NOAA") satellites and vegetation vigor indices from the U.S. Geological Survey ("USGS") to estimate crop condition. The USGS indices are based on the NOAA satellite remote sensor data. NASS compares the current year vigor indices with prior year images, using change detection postprocessing, to determine if crop development is lower, higher or the same as prior years. NASS has not yet been able to convert the vigor indices into specific crop yield data (NASS, 1998). The Foreign

Agricultural Service ("FAS") collects and reports statistical information on global crop conditions and production. This USDA agency has developed crop models, which combine satellite images and weather data to estimate yield, plant growth stage, soil moisture and winterkill. The FAS analyzes over 9,500 multispectral satellite images per year (FAS, 1998).

European Commission

The European Commission ("EC") through its complex networks of associations and departments, under the EC's research affiliate the Joint Research Centre ("JRC"), controls the research activity for country members of the European Union ("EU"). The JRC program, Monitoring Agriculture with Remote Sensing ("MARS"), provides the necessary technical support and image data to EU organizations such as the Directorate General for Agriculture and the European Statistical Office. Satellite data is used to measure crop acreage, type and yield. The resulting agricultural statistical information assists the EU in monitoring member states' compliance with EU agricultural policies.

12

History of Indian Remote Sensing Program

INTRODUCTION

India's remote sensing programme under the Indian Space Research Organization (ISRO) started off in 1988 with the IRS-1A, the first of the series of indigenous state-of-art operating remote sensing satellites,which was successfully launched into a polar sun-synchronous orbit on March 17, 1988 from the Soviet Cosmodrome at Baikonur.

It was a proud moment for the country and showed the maturity of the satellites in the various requirements for managing natural resources of the nation. It has sensors like LISS-I which had a spatial resolution of 72.5 meters with a swath of 148 km on ground. LISS-II had two separate imaging sensors, LISS-II A and LISS-II B, with spatial resolution of 36.25 meters each and mounted on the spacecraft in such a way to provide a composite swath of 146.98 km on ground.These tools quickly enabled India to map, monitor and manage its natural resources at various spatial resolutions. The operational availability of data products to the user organisations further strengthened the relevance of remote sensing applications and management in the country.

IRS System

Following the successful demonstration flights of Bhaskhar and Bhaskara-2 satellites launched in 1979 and 1981, respectively, India began to develop the indigenous Indian Remote Sensing (IRS) satellite program to support the national economy in the areas of agriculture, water resources, forestry and ecology, geology, water sheds, marine fisheries and coastal management.

Towards this end, India had established the National Natural Resources Management System (NNRMS) for which the Department of Space (DOS) is the nodal agency, providing operational remote sensing data services. Data from the IRS satellites is received and disseminated by several countries all over the world. With the advent of high-resolution satellites new applications in the areas of urban sprawl, infrastructure planning and other large scale applications for mapping have been initiated.

The IRS system is the largest constellation of remote sensing satellites for civilian use in operation today in the world, with 11 operational satellites. All these are placed in polar Sun-synchronous orbit and provide data in a variety of spatial, spectral and temporal resolutions. Indian Remote Sensing Programme completed its 25 years of successful operations on March 17, 2013.

IRS data applications

Data from Indian Remote Sensing satellites are used for various applications of resources survey and management under the National Natural Resources Management System (NNRMS). Following is the list of those applications:

- Space Based Inputs for Decentralized Planning (SIS-DP)
- National Urban Information System (NUIS)
- ISRO Disaster Management Support Programme (ISRO-DMSP)
- Biodiversity Characterizations at landscape level.
- Preharvest crop area and production estimation of major crops.
- Drought monitoring and assessment based on vegetation condition.

- Flood risk zone mapping and flood damage assessment.
- Hydro-geomorphological maps for locating underground water resources for drilling well.
- Irrigation command area status monitoring
- Snow-melt run-off estimates for planning water use in down stream projects
- Land use and land cover mapping
- Urban planning
- Forest survey
- Wetland mapping
- Environmental impact analysis
- Mineral Prospecting
- Coastal studies
- Integrated Mission for Sustainable Development (initiated in 1992) for generating locale-specific prescriptions for integrated land and water resources development in 174 districts.

Indian Space Programme

Despite being a developing economy with its attendant problems, India has effectively developed space technology and has applied it successfully for its rapid development and today is offering a variety of space services globally. During the formative decade of 1960s, space research was conducted by India mainly with the help of sounding rockets. The Indian Space Research Organisation (ISRO) was formed in 1969. Space research activities were provided additional fillip with the formation of the Space Commission and the Department of Space by the government of India in 1972. And, ISRO was brought under the Department of Space in the same year. In the history of the Indian space programme, 70s were the era of Experimentation during which experimental satellite programmes like Aryabhatta, Bhaskara, Rohini and Apple were conducted. The success of those programme, led to era of operationalisation in 80s during which operational satellite programmes like INSAT and IRS came into being. Today, INSAT and IRS are the major programmes of ISRO.

The most significant milestone of the Indian Space Programme during the year 2005-2006 was the successful launch of PSLV-C6. On 5 May 2005, the ninth flight of Polar Satellite Launch Vehicle (PSLV-C6) from Satish Dhawan Space Centre (SDSC) SHAR, Sriharikota successfully placed two satellites-the 1560 kg CARTOSTAR-1 and 42kg HAMSAT-into a predetermined polar Sun Synchronous Orbit (SSO). Coming after seven launch successes in a row, the success of PSLV-C6 further demonstrated the reliability of PSLV and its capability to place payloads weighing demonstrated the reliability of PSLV and its capability to place payloads weighing up to 1600 kg satellites into a 600 km high polar SSO. The successful launch of INSAT-4A, the heaviest and most powerful satellite built by India so far, on 22 December 2005 was the other major event of the year 2005-06. INSAT-4A is capable of providing Direct-To-Home (DTM) television broadcasting Services.)

Indian National Satellite System

The Indian National Satellite (INSAT) system is one of the largest domestic communication satellite systems in the Asia-Pacific region. In the 1980s, it initiated a major revolution in India's communications sector and sustained the same later. The satellites of INSAT system, which are in service today, are INSAT-2F, INSAT-3A, INSAT-3B, INSAT-3C, INSAT-3E, KALPANA-1, GSAT-2, EDUSAT and INSAT-4A, that was launched recently. The system provides a total of about 175 transponders in the C, Extended C and Ku-bands. Being a multipurpose satellite system, INSAT provides services to telecommunications, television broadcasting, weather forecasting, disaster warning and Search and Rescue fields.

INSAT system is also providing meteorological services through Very High Resolution Radiometer and CCD cameras on some of its spacecraft. This apart, cyclone monitoring through meteorological imaging and issue of warnings on impending cyclones through disaster warning receivers have been operationalised. For this, 350 receivers have been installed along the east and west coasts of India.

Indian Remote Sensing Satellite System

India has the largest constellation of Remote Sensing Satellites, which are providing services both at the national and global levels. From the Indian Remote Sensing (IRS) Satellites, data is available in a variety of spatial resolutions staring from 360 metres and highest resolution being 2.5 metres. Besides, the state-of-the-art cameras of IRS spacecraft take the pictures of the Earth in several spectral bands. In future, ISRO intends to launch IRS spacecraft with better spatial resolution and capable of imaging day and night. The satellites of IRS system which are in service today are IRS-1C, IRS-ID, IRS-P3, OCEANSAT-1, Technology Experimental Satellite (TES), RESOURCESAT-1, and the recently launched CARTOSAT-1 capable of taking stereo pictures. The upcoming Remote Sensing Satellite are Cartosat-2, RISAT (Redar Imaging Satellite) and Oceansat-2.

Launch Vehicles

After successfully testing the first indigenous launch vehicle SLV-3 in 1980, ISRO built the next generation Augmented Satellite Launch Vehicle (ASLV). ISRO's Launch Vehicle Programme had a giant leap with the successful launch of IRS-P2 spacecraft onboard the Polar Satellite Launch Vehicle (PSLV) in October 1994. On 18 April 2001, India successfully launched is Geosynchronous Satellite Launch Vehicle (GSLV). Technology development for advanced launch vehicles made good progress with the breakthrough achieved during the year in Supersonic Combustion Ramjet (SCRAMJET) to be employed in Air-Breathing engine. This is an important element in the launch vehicle technology development. Concepts for reusable launch vehicle are also being studied.

Polar Satellite Launch Vehicle

The four stage PSLV is capable of launching upto 1,600 kg satellites into a 620 km polar orbit. It has provision to launch payloads from 100 kg micro-satellites or mini or small satellites in different combinations. It can also launch one-two class payloads into Geosynchronous Transfer Orbit (GTO). So far, it has performed nine missions with eight consecutive successes. The latest launch of PSLV (PSLV-C6) was on 5 May 2005 during which the vehicle precisely placed the 1560 kg CARTOSAT-1 and the 42 kg HAMSAT into a 620 km high polar SSO.

Geosynchronous Satellite Launch Vehicle

The GSLV was successful on its very first test flight. After its successful second flight on 8 May 2003, it was commissioned. This was followed by the success of its third flight on 20 September 2004.

The GSLV is capable of launching 2,000 kg class satellites into Geosynchronous Transfer Orbit (GTO). The development of Indigenous cryogenic stage to be used as the third stage of GSLV made further progress during the year. The cryogenic engine which forms part of this stage, has already been successfully qualified. GSLV-Mk III, a new version of GSLV and capable of launching spacecraft weighing upto 4 tonnes to GTO is under development.

Launchinfrastructure

An elaborate launch infrastructure exists at the Satish Dhawan Space Centre (SDSC) SHAR, Sriharikota Island on the East Coast of India which is about 100 km from Chennai. Sriharikota is located at 13$dG North latitude. From here, satellites can be launched into a variety of orbital inclinations starting from 18$dG and extending upto 99$dG. Full-fledged facilities for satellite integration, assembly and launch exist there. Sriharikota also houses a Telemetry, Tracking and Command network for tracking satellites and monitoring them. The newly built Second Launch Pad at SDSE SHAR as a redundancy to the existing launch pad, and to cater to the requirement of GSLV-Mk III as well as other future launch vehicles, was commissioned on 5 May 2005 with the successful launch of PSLV-C6.

Space-industry Co-operation

One of the important features of the Indian Space Programme since its inception has been the co-operative approach with the Indian industries. The Department of Space (DOS) has established linkages with about 500 industries in small, medium and large-scale sectors, either through procurement contracts, know-how transfers or provision of technical consultancy. Because of its association with the space programme, the space industry is now capable of meeting the challenges in terms of adopting advanced technologies or handling complex manufacturing jobs.

Interface with Academic and R&D Institutions

The ISRO has an active programme to interact with academic and research institutions all over the country for the benefit of our space programme. In this regard, the Sponsored Research Programme (RESPOND) is an important component of DOS. Under RESPOND, DOS support research and educational activities at universities, individual colleges, and at the Indian Institutes of Technology as well as other research institutions. During the year 2005-2006, 13 projects were successfully completed and 62 new projects were initiated at 42 academic institutions comprising universities, colleges and research institutions. In addition to research projects, DOS supported 73 conferences, symposia, educational and promotional activities in the areas of importance to ISRO, besides providing support to ISRO-institutional chairs at reputed institutions.

DETERMINATION OF SCALE OF AERIAL PHOTOGRAPHS

Before a photograph can be used as a map supplement or substitute, it is necessary to know its scale. On a map, the scale is printed as a representative fraction that expresses the ratio of map distance to ground distance, RF = MD - GD. On a photograph, the scale is also expressed as a ratio, but is the ratio of the photo distance (PD) to ground distance, RF = PD - GD. The approximate scale or average scale (RF) of a vertical aerial photograph is determined by either of two methods; the comparison method or the focal length flight altitude method.

Small-scale and Large-scale aerial images

Small scale images, with a ratio of 1:25000 or 1:50000 for example, are those which cover a large area with less detail. A large scale image, around 1:3000 or 1:5000 for example, will cover a smaller area but will show ground features in more detail.

This photograph was taken in 1988 during the All Scotland Survey. It has a nominal scale of 1:24000, which means that every centimetre on the image represents 24,000 centimetres, or 240 metres on the ground.

This photograph shows the former docks area of Glasgow at a larger scale, 1:9800.

In this image, every centimetre on the image represents 9,800 centimetres, or 98 metres on the ground.

The area covered by this image represents a very small part of the lower-left corner of the 1988 photograph of Glasgow, above.

The buildings and landmarks surrounding the Glasgow docks area can be seen in much more detail in this larger-scale image.

DETERMINATION OF OBJECT HEIGHT ON AERIAL PHOTOGRAPHS

Aerial photography (or airborne imagery) is the taking of photographs from an aircraft or other flying object.

Platforms for aerial photography include fixed-wing aircraft, helicopters, unmanned aerial vehicles (UAVs or "drones"), balloons, blimps and dirigibles, rockets, pigeons, kites, parachutes, stand-alone telescoping and vehicle-mounted poles.

Mounted cameras may be triggered remotely or automatically; hand-held photographs may be taken by a photographer.

Aerial photography should not be confused with air-to-air photography, where one or more aircraft are used as chase planes that "chase" and photograph other aircraft in flight.

INTERPRETATION OF SINGLE VERTICAL AERIAL PHOTOGRAPHS

Aerial photographs are generally classified as being vertical or oblique. A vertical photograph is one which has been taken with the camera axis directed toward the ground as vertically as possible, while an oblique photograph is one which has been taken with the camera axis directed at an inclination to the ground.

Vertical

This image is a vertical aerial photograph of the River Forth at Throsk, Stirlingshire, in 1988.

Vertical photographs are the most common type of aerial photograph for remote sensing and aerial survey purposes. They can be scaled, allowing objects and distances to be measured, aiding in their identification.

When viewed in stereo, vertical photographs can give information about the height or the vertical characteristics of landmarks and buildings.

Vertical aerial photographs can provide very useful information, in conjunction with maps and other sources, when searching for unexploded ordnance or assessing property boundaries, for example.

INTERPRETATION OF STEREO PAIR OF AERIAL PHOTOGRAPHS

Consider the different shapes and what they might be. Is it a large or small scale image?

Does this knowledge change how you perceive certain shapes? Is the image taken directly from above or at an angle? Are there any shadows? Is there any variation in tone? How are different features related to each other?

These are the questions a photographic interpreter or imagery analyst would ask when studying an aerial image in order to extract useful information from it.

Shape

Shape refers to the general form, configuration or outline of individual objects and is normally the first thing the eye sees in an aerial photograph.

In many cases it provides immediate identification, however many objects have a similar shape. A circular object could be a gas holder, a chimney, a well or a roundabout, depending on its size.

Size

The size of a feature in an aerial photograph can be obtained in two ways:

- By comparing the feature to a recognised object - if the object is a well it would be smaller than a house, for example.
- By measurement - if the scale of the image is calculated, the exact size of the object can be determined.

Shadow

Shadows can reveal the profile or outline of an object, and are useful for determining its shape, height and construction.

- shadows can indicate whether a feature is of open or solid construction
- shadows can reveal the shape of a roof structure
- low-angle sunlight will cast long shadows from the slightest ground features, such as archaeological remains

- tree shadows can be used to distinguish between coniferous and deciduous types
- the height of an object can be calculated from its shadow length, if the time and location of imaging is known

Tone

The relative brightness of objects in an image is described by the term 'tone'. Tone is related to texture and colour and is the measure of the amount of light reflected back to the camera by an object. The colour of an object has less effect on tone than is generally expected.

The smoother the surface of an object, the greater is the amount of light reflected back to the camera - this is why a black asphalt road may appear lighter than a green field of grass. Note the difference in tone between the canals, streets and houses in this image of Amsterdam.

In images of arable land, subtle tonal variations can highlight differences in drainage and vegetation.

Association

Objects can often be identified by looking at their surroundings. In the identification of the circular objects referred to in 'Shape', above, a gas holder would be sited adjacent to a gasworks, a chimney would be sited within a factory and a roundabout would be situated at a road junction.

In the identification of military features in the field, the main associated features are tracks, emplacements and trenches.

Stereo Photography

Stereo photography is a means by which 3-dimensional features such as terrain depth and the relative heights of buildings can be perceived by combining views from one or more 2-dimensional photos.

As strange as this may sound, it actually works by exactly the same principle that enables human eyesight to have depth perception.

When a pair of eyes focuses on an object, each eye views the object from a slightly different angle.

When the brain combines both of these images by automatically comparing and calculating the difference between the two perspectives, the two angles are "merged together", resulting in the ability to determine that one detail is closer to the viewer than another.

Much the same process happens when an airplane takes aerial photography. As the plane passes over the ground, the camera takes pictures at certain intervals.

When one picture is taken, the features in the center of the photo appear in a vertical, directly "top down" perspective, while all the other features around the center appear at increasingly greater angles as they near the outer edges of the frame. By the time the very next photo is taken, the plane has moved forward and the new picture displays a new view with different perspectives from the one before.

The camera shots and the speed of the plane are regulated to provide a mutually overlapping portion on both photos.

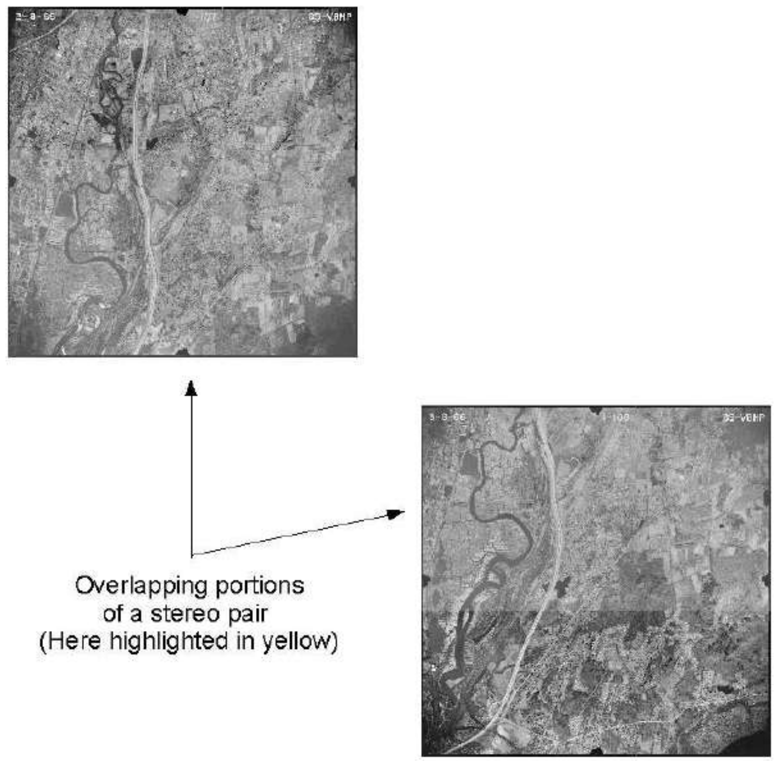
Within this portion (called "stereo overlap") all of the visible features, therefore, appear at different perspectives between one photo and the other. A pair of photos with stereo overlap is called a "stereo pair".

Often, a single photo of a stereo pair will be referred to as the "stereo mate" of the other in the pair.

This image (above) shows how the visible angles of landscape features, such as the buildings, appear to increase the farther away they are from the center of the photo. This is highlighted in this second view of the same image (below). Notice how the angles of the buildings (some of which are now highlighted in yellow for better clarity) appear to stretch off in the general directions of the red arrows. The buildings closer to the edge show the greatest angle increase.

Example Stereo Pair

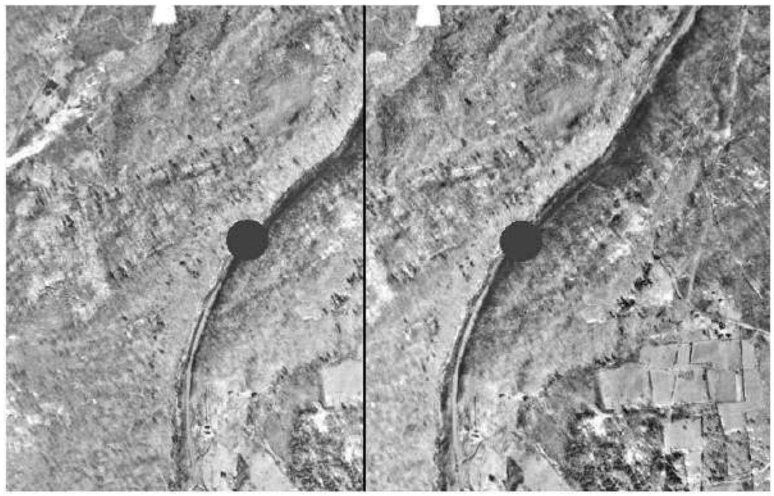
Both photos have an amount of surface area in common.



(Notice that these particular photos were taken along a north-south flight path. The overlapping portion of the northernmost frame comprises the bottom 60% of the entire photo, while the matching portion of its stereo mate—the southernmost frame—comprises the top 60% of the photo.)

If these different perspectives are combined, the result is a view that looks 3-dimensional, in just the same way that the different perspectives of two eyes allow depth perception when combined.

The best method to view the stereo overlap of a stereo pair is with a device, such as a tabletop stereo viewer, that is designed to merge the landscape features when viewed through it. However, it is possible to see the effect with a stereo pair of photos simply by crossing your eyes, as with the following example image.



This image is composed of two matching portions from a stereo pair laid side by side. The red dot in the center of each portion has been added here as a guide for demonstration purposes. If you begin to cross your eyes when looking at this picture, you should see that the two red dots seem to spilt into four dots. If you can cross your eyes further, you should notice that the two middle dots of the four will move closer to one another. Once they are lined up over top of each other, the area around them should look as though it has depth to it (despite the fact that you may also perceive the appearance of a certain amount of "double exposure"). Even without a guide such as the red dots, one need only find a distinctive feature present in both photos within the stereo overlap (such as a road, building, stream, etc.) and join them in the same manner.

INTRODUCTION TO REFERENCE SYSTEM OF IRS SATELLITES

Following the successful demonstration flights of Bhaskara-1 and Bhaskara-2 - experimental Earth observation satellites developed and built by ISRO (Indian Space Research Organization) - and launched in 1979 and 1981, respectively, India began the development of an indigenous IRS (Indian Remote Sensing Satellite) program. India realized quite early that sustaining its space program in the long run would depend on indigenous technological capabilities (in particular, US export restrictions made this clear). Keeping this in mind, besides building satellites, India embarked as well on satellite launch vehicle development in the early 1970s. As a consequence, India has two very capable launch systems at the start of the 21st century, namely PSLV (Polar Satellite Launch Vehicle) and GSLV (Geosynchronous Satellite Launch Vehicle).

IRS is the integrated LEO (Low Earth Orbit) element of India's NNRMS (National Natural Resources Management System) with the objective to provide a long-term spaceborne operational capability to India for the observation and management of the country's natural resources (applications in agriculture, hydrology, geology, drought and flood monitoring, marine studies, snow studies, and land use). The intend of the program is to create an environment of new perspectives for the Indian research community as a whole, to stimulate the development of new technologies and applications, and to utilize the Earth resources in more meaningful ways.

Note: The INSAT system is India's GEO (Geosynchronous Earth Orbit) element, providing for simultaneous domestic communications and earth observation functions.

The IRS program started in the mid 1980s. Eventually, a continuous supply of synoptic, repetitive, multispectral data of the Earth's land surfaces was obtained (similar to the US Landsat program). In 1995, IRS imagery was made available to a larger international community on a commercial basis. The initial program of Earth-surface imaging was extended by the addition of sensors for complementary environmental applications. This started with the IRS-P3 satellite which is flying MOS (Multispectral Optoelectronic Scanner) for the measurement of ocean color. The IRS-P4 mission is dedicated to ocean monitoring.

Note: The availability of Landsat imagery created a lot of interest in the science community. The Hyderabad ground station started receiving Landsat data on a regular basis in 1978. The Landsat program with its design and potentials was certainly a great model and yardstick for the IRS program.

- The first generation satellites IRS-1A and 1B were designed, developed and launched successfully during 1988 and 1991 with multispectral cameras with spatial resolutions of 72.5 m and 36 m, respectively. These early satellites were launched by Russian Vostok boosters from the Baikonur Cosmodrome.
- Subsequently, the second generation remote sensing satellites IRS-1C and -1D with improved spatial resolutions have been developed and successfully launched in 1995 and 1997, respectively. IRS-1C/1D data has been used for cartographic and town planning applications.

IRS-1A (Indian Remote Sensing Satellite-1A)

The spacecraft bus is box-shaped (1.6 m x 1.6 m x 1.5 m) with two solar panels (8.6 m^2). The S/C structure is made of aluminum/aluminum honeycomb. The satellite is three-axis stabilized utilizing a zero momentum system. Hydrazine thrusters (80 kg fuel) are also used for control and momentum dumping. The IRS series satellites are built around a zero-momentum reaction wheel based system.

Gyro-based attitude reference using quaternion propagation with attitude updates from Earth sensors and sun sensors (CCD-based is used for yaw angle measurements) provide the high-pointing accuracy and stability required for the imaging payload. Attitude is sensed by Earth sensor, sun sensor, star sensor and dynamically tuned gyros. The actuators are reaction wheels (4), magnetic torquers, and hydrazine thrusters (sixteen 1-newton thrusters). A pointing accuracy of 0.3º is achieved in pitch/roll and 0.5º in yaw. Attitude determination accuracy of ±0.1º. Total S/C mass = 975 kg (at launch), power = 700 W, two NiCd batteries (40 Ah) provide power for the eclipse phase of the orbit. The design life is three years.

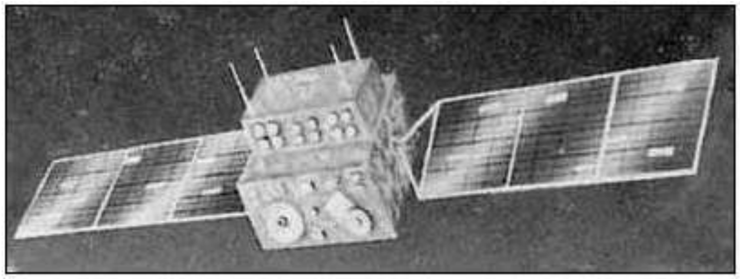


Figure : Illustration of the IRS-1A spacecraft (image credit: ISRO)

Launch: March 17, 1988 on a Russian launch vehicle Vostok-2M from the Baikonur Cosmodrome, Kazhakstan.

Orbit: sun-synchronous orbit, nominal altitude = 904 km, inclination = 99.049º, period = 103.2 minutes; the repeat cycle = 22 days; equator crossing at 10:26 hours on descending node.

Application: Land use, agriculture, forestry, hydrology, soil classification, coastal wetland mapping, natural resources (in particular pinpointing likely groundwater locations), disaster monitoring, cartography, etc.

RF communications: Payload downlink in S-band and X-band. The satellite carried a real time LISS-2A/B data downlink in X-band with a transfer rate of 10.4 Mbit/s (2 links). The data was downlinked to a 10 m dish antenna at Shadnagar on 20 W. The LISS-I data downlink was on 5 W S-band at a rate of 5.4 Mbit/s. There was no onboard recorder. The payload data were PCM/BPSK modulated.

The IRSO S/C control center is in Bangalore. TT&C function is provided by ISTRAC (ISRO Tracking Network), supported by DLR (GSOC, Weilheim), NOAA (Fairbanks), ESA (Malindi) and the USSR (Bearslake) ground stations.

DATA PRODUCTS AND FORMATS

IRS Data Products

Earth Observation (EO) from satellite platforms is proved to be an indispensable tool for natural resources mapping, monitoring and management, including environmental assessment at global, regional and local levels. This is particularly due to multi-platform, multi-resolution, multi-temporal and synoptic viewing capabilities from space platforms. Remote Sensing Satellites were commissioned as an integrated facility with Data acquisition and Product generation at Shadnagar complex of National Remote Sensing Centre (NRSC), formerly known as National Remote Sensing Agency (NRSA), for Landsat MSS data reception, way back in 1979. With the availability of data from the operational Indian Remote Sensing Satellites (IRS), from 1988, NRSA started disseminating IRS satellite data to the user community.

Over a period of time, the Indian EO programme has evolved into a fully operational system providing rich data services to the user community. Data from EO satellites support wide range of information needs of the user community for better understanding of the Earth system at global to local scale and helps in providing information on natural resources, agricultural, water, land use, forests, weather, natural disasters and so on. Today, India is one of the leading countries in the world for the operational use of satellite remote sensing for National development. The Indian EO System is widely acclaimed around the world for its application driven approach. The system operates under a unique national level institutional framework, National Natural Resources Management System (NNRMS) established under the aegis of the Planning Commission.

With a host of payloads onboard the series of Indian Remote Sensing (IRS) Satellites, the Indian EO system is providing varieties of operational services to the user community in the country. The Planning Committee of NNRMS (PC-NNRMS), chaired by Member (Science), Planning Commission, oversees various aspects of EO technology utilization in the country and provides necessary guidance for implementation of application programmes of national importance. The nine high-power Standing Committees, constituted under NNRMS, provides the impetus for operational use of remote sensing and allied technologies in different thematic areas, through respective Ministries. Government has adopted a comprehensive Remote Sensing Data Policy (RSDP) for the acquisition and distribution of satellite remote sensing data - from Indian and foreign satellites for civilian users in India.

The Policy comprehensively covers guidelines for satellite data acquisition and distribution in the country and also for licensing the IRS capacities to other countries. Department of Space is the nodal agency for implementing the Policy. In particular, the Policy streamlines the distribution of high-resolution data to Government users; private users involved in developmental activities with government and other private/ academic/ foreign users. The RSDP is envisaged to be a step towards making transparent procedures for satellite data distribution, including those from high resolution imaging systems, and, without being restrictive. This has helped in regulating the process of data distribution, so that the Indian users are not denied access to this valuable information - which has become a main¬stay of the natural resources management activities of the country. Through this policy, the National Remote Sensing Centre (NRSC) of the Indian Space Research Organization (ISRO)/ DOS is vested with the authority to acquire and disseminate all satellite remote sensing data in India, both from Indian and foreign satellites.

The NRSC has wealth of images from Indian and foreign remote sensing satellites in its archives and also has the capability to acquire data pertaining to any part of the globe on demand. NRSC also supports, through ANTRIX, establishment of International Ground Stations and International reseller network to receive, process and market IRS data products globally.

NRSC provides end-to-end solutions for utilization of data for natural resources management, geospatial applications and information services. NRSC facilitates several remote sensing & GIS application projects for natural resources and environmental management catering to food security, water security, energy security and sustainable development. NRSC is also providing single window, disaster management support services through the Decision Support Centre. NRSC also implemented Bhuvan, a Geoportal of Indian Space Research Organization, showcasing Indian Imaging Capabilities in multi-sensor, multi-platform and multi-temporal domain. This Earth browser gives a gateway to explore and discover virtual earth in 3D space with specific emphasis on Indian Region, thematic data dissemination, Pocket Bhuvan and EO data down loads.

The satellite data are acquired, processed and disseminated to the user community as standard/ value added products to the Indian and International user community. In order to minimize the turn-around-time from data acquisition to product delivery and to provide near real time data for critical applications, Integrated Multi-mission Ground Segment for Earth Observation Satellites (IMGEOS) has been conceived and is being implemented at Shadnagar Complex of NRSC.

Data Format in Remote

Remote sensing data or image data is a digital picture or representation of various objects on the Earth's surface. The picture is a systematic arrangement of raster cells. Each of the raster cells, depending on the intensity of radiation received, contains a digital number between a certain range, for example, 0- 127 (7 bit image) or 0-255 (8 bit image) and so on, depending upon radiometric processing capacity of the detector system of the sensor. Each number (of each cell) in an image file is a data file value, sometimes also called pixel (abbreviation of picture element), and data file value is the measured brightness value of the pixel at a specific wavelength.

Raster image data are laid out in a grid format similar to squares on a checkerboard. These raster cells are assigned gray shades from darkest shade for zero digital number to the brightest white shade for digital number 127 or 255 or 511 and so on, and comparative grades of dark and white shades are assigned in between from digital numbers 1-126 or 1-254 or 1-510 and so on. Image data format can be defined as the sequential arrangement of pixels, representing a digital image in a computer compatible storage medium, such as a compact disk (CDs/DVDs).

Superposition of any three bands of data, each of which is developed in blue, green and red shades gives a color composite image of the area. That means, remote sensing image data, stored in data files/image files on magnetic tapes, compact disks (CDs/ DVDs) or other media, consist only of digital numbers.

These representations of numbers form the B & W or color images when they are displayed on a screen or output on a hard copy. Thus, the image has to be retained in its digital form in order to carry computer processing/ classification. The digital output is supplied on a suitable computer compatible storage media, such as DVDs, CD-ROMs, DAT, etc., depending on user requests. The data may be arranged in band sequential (BSQ), band interleaved by line (BIL) or band interleaved pixel (BIP) formats. Similarly, the concept of image data format comes in, with the question of how to arrange these pixels to achieve optimum level of desired processing and display. Let us look at the following example, a data file in jpg format is a compressed file in a small size, say 10MB; whereas, the same file in tiff format is uncompressed and its size can go up to 100MB. What happens in these two cases of files is the data transfer is easier with small size file, like a jpg file than in tiff format.

INTERPRETATION OF SATELLITE IMAGES

Satellite images are like maps: they are full of useful and interesting information, provided you have a key. They can show us how much a city has changed, how well our crops are growing, where a fire is burning, or when a storm is coming. To unlock the rich information in a satellite image, you need to:

1. Look for a scale
2. Look for patterns, shapes, and textures
3. Define the colors (including shadows)
4. Find north

5. Consider your prior knowledge

These tips come from the Earth Observatory's writers and visualizers, who use them to interpret images daily. They will help you get oriented enough to pull valuable information out of satellite images.

Look for a Scale

One of the first things people want to do when they look at a satellite image is identify the places that are familiar to them: their home, school, or place of business; a favorite park or tourist attraction; or a natural feature like a lake, river, or mountain ridge. Some images from military or commercial satellites are detailed enough to show many of these things. Such satellites zoom in on small areas to collect fine details down to the scale of individual houses or cars. In the process, they usually sacrifice the big picture.

NASA satellites take the opposite approach. Earth science researchers typically want a wide-angle lens to see whole ecosystems or atmospheric fronts. As a result, NASA images are less detailed but cover a wider area, ranging from the landscape scale (185 kilometers across) to an entire hemisphere. The level of detail depends on the satellite's spatial resolution. Like digital photographs, satellite images are made up of little dots called pixels. The width of each pixel is the satellite's spatial resolution.

Commercial satellites have a spatial resolution down to 50 centimeters per pixel. The most detailed NASA images show 10 meters in each pixel. Geostationary weather satellites, which observe a whole hemisphere at a time, are much less detailed, seeing one to four kilometers in a pixel.

Depending on the image resolution, a city may fill an entire satellite image with grids of streets or it may be a mere dot on a landscape. Before you begin to interpret an image, it helps to know what the scale is. Does the image cover 1 kilometer or 100? What level of detail is shown? Images published on the Earth Observatory include a scale.

You can learn different things at each scale. For example, when tracking a flood, a detailed, high-resolution view will show which homes and businesses are surrounded by water. The wider landscape view shows which parts of the county or metropolitan area are flooded and perhaps where the water is coming from. A broader view would show the entire region—the flooded river system or the mountain ranges and valleys that control the flow. A hemispheric view would show the movement of weather systems connected to the floods.

Look for patterns, shapes, and textures

If you have ever spent an afternoon identifying animals and other shapes in the clouds, you'll know that humans are very good at finding patterns. This skill is useful in interpreting satellite imagery because distinctive patterns can be matched to external maps to identify key features.

Bodies of water—rivers, lakes, and oceans—are often the simplest features to identify because they tend to have unique shapes and they show up on maps.

Other obvious patterns come from the way people use the land. Farms usually have geometric shapes—circles or rectangles—that stand out against the more random patterns seen in nature. When people cut down a forest, the clearing is often square or has a series of herring-bone linesthat form along roads. A straight line anywhere in an image is almost certainly human-made, and may be a road, a canal, or some kind of boundary made visible by land use.

Geology shapes the landscape in ways that are often easier to see in a satellite image. Volcanoes and craters are circular, and mountain ranges tend to run in long, sometimes wavy lines. Geologic features create visible textures. Canyons are squiggly lines framed by shadows. Mountains look like wrinkles or bumps.

These features can also affect clouds by influencing the flow of air in the atmosphere. Mountains force air up, where it cools and forms clouds. Islands create turbulence that results in swirling vortices or wakes in the clouds. When you see a line of clouds or vortices, they provide a clue about the topography of the land below.

Occasionally, shadows can make it hard to tell the difference between mountains and canyons. This optical illusion is called relief inversion. It happens because most of us expect an image to be lit from the top left corner. When the sunlight comes from another angle (especially from the lower edge), the shadows fall in ways we don't expect and our brains turn valleys into mountains to compensate. The problem is usually resolved by rotating the image so the light appears to come from the top of the image.

Define Colors

The colors in an image will depend on what kind of light the satellite instrument measured. True-color images use visible light—red, green and blue wavelengths—so the colors are similar to what a person would see from space. False-color images incorporate infrared light and may take on unexpected colors.

Water

Water absorbs light, so it is usually black or dark blue. Sediment reflects light and colors the water. When suspended sand or mud is dense, the water looks brown. As the sediment disperses, the water's color changes to green and then blue. Shallow waters with sandy bottoms can lead to a similar effect.

Sunlight reflecting off the surface of the water makes the water look gray, silver, or white. This phenomenon, known as sunglint, can highlight wave features or oil slicks, but it also masks the presence of sediment or phytoplankton.

Frozen water—snow and ice—is white, gray, and sometimes slightly blue. Dirt or glacial debriscan give snow and ice a tan color.

Plants

Plants come in different shades of green, and those differences show up in the true-color view from space. Grasslands tend to be pale green, while forests are very dark green. Land used for agriculture is often much brighter in tone than natural vegetation.

In some locations (high and mid latitudes), plant color depends on the season. Spring vegetation tends to be paler than dense summer vegetation. Fall vegetation can be red, orange, yellow, and tan; leafless and withered winter vegetation is brown. For these reasons, it is helpful to know when the image was collected.

In the oceans, floating plants—phytoplankton—can color the water in a wide variety of blues and greens. Submerged vegetation like kelp forests can provided a shadowy black or brown hue to coastal water.

Bibliography

Allen J.: *Cultural Industries and the Production of Culture,* Taylor & Francis, 2007.

Baker, J. A. : *Liberating Life: Contemporary Approaches to Ecological Theology,* Maryknoll, NY: Orbis Books, 1990.

Barnes, Thomas G. : *Origin and Destiny of the Earth's Magnetic Field,* Creation Research Society, London, 1983.

Carey, S. Warren : *Theories of the Earth and Universe*, Stanford, Stanford Univ. Press, 1988.

Carole L. : *Toward a Locational Definition of State Systems of Settlement,* USA, 1976.

Crumley, Carole L. : *Toward a Locational Definition of State Systems of Settlement*, USA, 1976.

Dicken, P. : *Wheels of Change, the Automobile Industry*, *Global shift, Mapping the Changing Contours of the World Economy*, London, 2007.

Earney, F.C.F.: *Marine Mineral Resources: Ocean Management and Policy*, London, Routledge, 1990.

Edwards, Denis : *Earth Revealing Earth Healing: Ecology and Christian Theology,* Collegeville: Liturgical Press, 2001.

Goodchild, M.F and Kemp, K.K: *National Centre for Geographic Information and Anlysis*, Univ. of California, USA, 1990.

Hall, Dorothy K., and Jaroslav Martinec : *Remote Sensing of Ice and Snow*, London, New York, Chapman and Hall, 1985.

Harvey, Francis : *A Primer of GIS, Fundamental Geographic and Cartographic Concepts,* The Guilford Press, 2008.

Juppenlatz, Morris : *Geographic Information Systems and Remote Sensing*, McGraw Hill, New York, 1996.

Khan, M.M.A. : *Dictionary of Remote Sensing,* Manak, Delhi, 1993.

Kumar Yadav, Surendra : *Soil Ecology : Remote Sensing and Geographic Information System in Management,* APH, Delhi, 2007.

Lindgren, David T. : *Land Use Planning and Remote Sensing*, Canada, Kluwer Academic Pub., 1985.

Mack, Pamela Etter : *Viewing the Earth: The Social Construction of the Landsat Satellite System*, Cambridge, MIT Press, 1900.

Matthews, M. C. *: Remote Sensing in Civil Engineering*, Glasgow, Surrey University Press; New York, 1985.

McDonnell, R.A. : *Principles of Geographical Information Systems*, Oxford University Press, Oxford, 1998.

Mukherjee, Saumitra : *Remote Sensing Applications in Applied Geosciences,* Manak, Delhi, 1999.

Nierenberg, W. A.: *Encyclopedia of Earth System Science*, San Diego, California: Academic Press, 1992.

Okoshi, T. : *Three-dimensional Imaging Techniques*, Academic Press Inc., New York, 1976.

Oldroyd, David : *Earth Cycles: A Historical Perspective*, Westport, Connicticut: Greenwood Press, 2006.

Ramasamy, S. : *Remote Sensing in Water Resources,* Rawat, 2005.

Raymond M. : *Laser Remote Sensing: Fundamentals and Applications*, Malabar, Fla., Krieger Pub. 1992.

Richards, John A. : *Remote Sensing Digital Image Analysis: An Introduction*, Berlin, New York, 1993.

Ruth Towse : *Developments in the Economics of Copyright, Research and Analysis,* Edward Elgar Publishing, Cheltenham, United Kingdom, 2005.

Scott, Allen J.: *Cultural Industries and the Production of Culture,* Taylor & Francis, 2007.

Sebba, J.: *Geography for All*, London, David Fulton, 1996.

Singh, Surendra : *Remote Sensing of Indian Deserts,* Manak, Delhi, 2008.

Sokhi, B.S. and S.M. Rashid : *Remote Sensing of Urban Environment,* Manak, Delhi, 1999.

Steven J.: *Ecopreneuring: The Complete Guide to Small Business Opportunities from the Environmental Revolution*, John Wiley and Sons, 1992.

Thomashow, Mitchell: *Ecological Identity: Becoming a Reflective Environmentalist,* Cambridge, MIT Press, 1995.

Walter Leal: *Innovation, Education, and Communication for Sustainable Development*, New York, Peter Lang Publishing, 2006.

Wendy J. Gordon : *The Economics of Copyright, Developments in Research and Analysis* Edward Elgar, 2003.

Index

L

M

N

P

R

S

T

❑❑❑