Chapter 2

Study of Feature Extraction Methods and a Survey of Existing Works

2.1 Introduction
The main subject of this thesis lies in the overlap area of two different research fields:
1. Remote Sensing: Input data are remotely sensed images.
2. Feature Extraction: Remotely sensed data are digital images which are the basis for feature extraction methods and algorithms leading to meaningful image objects.

This chapter presents a brief review of these research fields relevant to the thesis.

2.2 Remote Sensing
According to Lillesand, T. & Kiefer, R. W. (1987), remote sensing is the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, phenomenon under investigation. Although this is a rather broad definition of remote sensing, this work is restricted to digital images that are acquired by sensors mounted on satellites.
2.2.1 History
An essential element in remote sensing is the ability to record images of the earth’s surface. In 1858, photography was used to acquire an aerial view of the earth’s surface from a captive balloon. In the following years, technology and methods leading to aerial images of the earth improved steadily. The next step was the use of air-planes for taking aerial images. Their manoeuvrability is an essential prerequisite for systematic coverage of a region by aerial photography.

The developments in space technology and the first launch of a meteorological satellite (TIROS-1) in 1960 provided the basis for earth observation satellites (Campbell, J. B., 1987). Landsat 1, launched in 1972, was the first satellite system dedicated to civilian land observation applications. Several other missions followed this successful satellite. With the Landsat missions it was possible to routinely acquire multispectral remote sensing imagery of large areas. In the 80’s Landsat TM and the French SPOT programme have developed further (Baudoin, A., 1995), and a great number of applications were based on these satellite data. Starting from 1990, the abolishment of military restrictions led to a commercialisation of the remote sensing field. In addition, new sensors with higher spatial resolution were developed which increased the research possibilities in this area drastically.

2.2.2 System Overview of Remote Sensing
Remote sensing image acquisition can be seen as part of a process cycle as shown in fig. 2.1. Remote sensing data of the earth surface are obtained by sensors mounted on aircrafts or satellites. This leads to either analog or digital data that have to be processed according to the application. The following list gives a glimpse of tasks necessary for the analysis and processing of raw data:

- Georeferencing, i.e. to register the image into a map coordinate system.

- Radiometric correction, i.e. to correct for atmospheric and surface topography influences.

- Classification, i.e. to assign land cover classes to the image.
It is understood that these tasks require reference data such as control points for georeferencing, digital elevation models (DEM), training data sets, etc. The final products are maps and/or lists of objects, i.e. a description of the current status of the observed part of the earth surface. These results are incorporated to a geographical information system (GIS). The analysed data are made available to users, e.g. companies or policymakers. These users take measures based upon the received information, e.g. nature protection measures, road construction, forest development and cultivation measures, etc., which have in turn an influence on the earth surface.

2.2.3 Radiation Paths and Principles
Remote sensors detect electromagnetic energy that various earth surface features emit and reflect. The sources for the energy are the sun or the earth itself. The sun radiation propagates through the atmosphere, is reflected by the earth surface and retransmitted through the atmosphere. The earth itself also emits radiation (in the thermal or far-infrared) that can be detected by sensors. Sensor systems that record this kind of radiation are called passive systems since they detect only naturally occurring radiation. Systems that actively emit radiation and detect the reflected
radiation are called active systems (e.g. RADAR, LIDAR), which is indicated in fig. 2.2.

Figure 2.2: Passive & Active Sensors

2.2.4 Optical Satellite Data

All passive optical sensors (visible through thermal infrared spectral range) operate on the same principles of optical radiation transfer, image formation, and photon detection. The details of sensor construction and detector materials vary with wavelength range, resolution, etc. These sensors supply the data as digital multispectral images. A digital satellite image is characterised by spatial, radiometric and spectral resolution (Schowengerdt, R. A., 1997).

2.2.4.1 Spatial Resolution

The detail discernible in an image is dependent on the spatial resolution of the sensor and refers to the size of the smallest possible feature that can be detected. Spatial resolution of passive sensors depends primarily on their Instantaneous Field of View (IFOV). The IFOV is the angular cone of visibility of the sensor (A) and determines the area on the Earth’s surface which is "seen" from a given altitude at one particular moment in time (B). The size of the area viewed is determined by multiplying the IFOV by the distance from the ground to the sensor (C).

Figure 2.3: Spatial Resolution of a Sensor
2. Study of Feature Extraction Methods and a Survey of Existing Works

This area on the ground is called the resolution cell as shown in fig. 2.3 and determines a sensor's maximum spatial resolution. If a sensor has a spatial resolution of 20 metres and an image from that sensor is displayed at full resolution, each pixel represents an area of 20m x 20m on the ground. Images where only large features are visible are said to have coarse or low resolution. In fine or high resolution images, small objects can be detected. Commercial satellites provide imagery with resolutions varying from a few metres to several kilometres. Generally speaking, the finer the resolution, the less total ground area can be seen.

2.2.4.2 Spectral Resolution
Spectral resolution describes the ability of a sensor to define fine wavelength intervals. The finer the spectral resolution, the narrower the wavelength ranges for a particular channel or band. The simplest form of spectral resolution is a sensor with one band only, which senses visible light. An image from this sensor would be similar in appearance to a black and white photograph from an aircraft. A sensor with three spectral bands in the visible region of the electro-magnetic spectrum would collect similar information to that of the human vision system. Multispectral and hyper spectral images consist of several bands of data (fig. 2.4). For visual display, each band of the image may be displayed one band at a time as a grey scale image, or in combination of three bands at a time as a color composite image.

2.2.4.3 Radiometric Resolution
The radiometric resolution of an imaging system describes its ability to discriminate very slight differences in energy. The finer the radiometric resolution of a sensor, the more sensitive it is to detecting small differences in reflected or emitted energy.
Imagery data are represented by positive digital numbers which vary from 0 to $2^{(n-1)}$, where ‘n’ is the number of bits used for coding data in binary format. Each bit records an exponent of power 2 (e.g. 1 bit=$2^1=2$). The maximum number of brightness levels available depends on the number of bits used in representing the energy recorded. If a sensor used has 8 bits to record the data, there would be $2^8=256$ digital values available, ranging from 0 to 255. However, if only 4 bits were used, then only $2^4=16$ values ranging from 0 to 15 would be available. Thus, the radiometric resolution is same as the gray level resolution of the image. By comparing a 2-bit image with an 8-bit image, we can see that there is a large difference in the level of detail discernible depending on their radiometric resolution which is shown in fig.2.5.

![Figure 2.5: (a) Poor Radiometric Resolution (2-bit image) (b) Good Radiometric Resolution (8-bit image)](image)

### 2.2.4.4 Temporal Resolution

Temporal resolution is the time period between successive image acquisitions of a given area. Spectral characteristics of features may change over time and these changes can be detected by collecting and comparing multi-temporal imagery. By imaging on a continuing basis at different times we are able to monitor the changes that take place on the Earth's surface, whether they are naturally occurring (such as changes in natural vegetation cover or flooding) or induced by humans (such as urban development or deforestation).
2.2.5 Processed and Unprocessed Satellite Images

Any remotely sensed image, regardless of whether it is acquired by a multispectral scanner on board a satellite, a photographic system in an aircraft, or any other platform/sensor combination, will have various geometric distortions. This problem is inherent in remote sensing, as we attempt to accurately represent the three-dimensional surface of the Earth as a two-dimensional image. All remote sensing images are subject to some form of geometric distortions, depending on the manner in which the data are acquired. These errors may be due to a variety of factors, including one or more of the following, to name only a few:

- the perspective of the sensor optics,
- the motion of the scanning system,
- the motion and (in)stability of the platform,
- the platform altitude, attitude, and velocity,
- the terrain relief, and
- the curvature and rotation of the Earth.

Prior to data analysis, initial processing on the raw data is usually carried out to correct for any distortion due to the characteristics of the imaging system and imaging conditions. Depending on the user's requirement, some standard correction procedures may be carried out by the ground station operators before the data is delivered to the end-user. These procedures include radiometric correction to correct for uneven sensor response over the whole image and geometric correction to correct for geometric distortion due to Earth's rotation and other imaging conditions.

Since most remotely sensed data require similar amounts of basic processing before they are usable, image distribution agencies have adopted a common set of processing "levels" to describe the types of processing done to their images before they send them out the door (Piwowar, J. M., 2001).

2.2.5.1 Level 0 Imagery

Level 0 imagery is raw instrument data, just as they were collected at the sensor. Since there are some fundamental corrections that should be applied to the data before they are usable, most agencies will not distribute level 0 imagery.
2.2.5.2 Level 1A Imagery
The next step up from level 0 is level 1A. Level 1A data have been corrected for detector variations referred to as radiometric correction. The number of detectors can vary from 16 for the Landsat Thematic Mapper to 6000 for the SPOT sensor. The Level 1A processing involves applying inter-detector equalizations through a series of data calibrations involving pre-launch values and in-flight measurements. Also, absolute calibration coefficients are posted in the ancillary data and can be used to convert the pixel values into real irradiance measurements.

2.2.5.3 Level 1B Imagery
Another type of systematic error inherent in space-borne data involves the geometry of the image product. The sensor is moving in orbit above the Earth's surface; the Earth itself is rotating beneath the satellite; and the sensor's view of the surface can be considerably more oblique towards the edges of the scene than it is directly below. All of these variations create distortions in the geometry of the imagery such as miss-aligned scan lines and non-uniform pixel sizes. Since all of these distortions are predictable and measurable, systematic corrections can be applied to the imagery to improve its geometric qualities. This is the operation of level 1B processing.

2.2.5.4 Level 2A Imagery
Level 2A images have been systematically mapped into a standard cartographic map projection based on a prediction of where the satellite was when the image was acquired. Therefore, level 2A images are frequently labelled as ‘geo-referenced’.

2.2.5.5 Level 2B Imagery
In this level, through a process called geometric correction or image rectification, the image analyst ‘registers’ the image to an existing base map by selecting pairs of well-defined points known as Ground Control Points (GCPs) from both the image and the map. When sufficient number of GCPs has been accurately identified the image can be geo-referenced so that its geometry will match that of the base map to which it was registered. The position accuracy of level 2B images generally matches the spatial resolution of the original data, except in areas of high local relief.

2.2.5.6 Level 3A Imagery
If the region is mountainous, we need to account for relief displacement in order to obtain consistently high position accuracies. In addition to manually locating ground
control points as in level 2B, Digital Elevation Model (DEM) is used for the relief displacement at differing elevations. This process is generically called ortho-rectification. The position accuracy of level 3A images matches the spatial resolution of the original data, including areas of high local relief.

2.2.5.7 Level 3B Imagery
Level 3B images have all the same attributes as level 3A images, but cover a larger area by mosaicking several scenes together. If we have access to a remote sensing image analysis system or robust GIS with geometric correction capabilities, then level 1B data are the most preferable one. Level 1B imagery has both inter-detector equalizations and systematic geometry corrections applied. It forms a good base for most remote sensing applications. If we don't have the necessary software and/or time to process the imagery, then we have to purchase the data at level 2B or 3A, depending on the amount of relief in the region. These images will import directly into a GIS and be ready for use.

2.2.6 Types of Satellite Images
Different satellites have different type of sensors which provide different types of images of the same region. Depending on the types of satellite images available there are mainly four categories:

2.2.6.1 Panchromatic Images (PAN)
An image collected in the broad visual wavelength range but rendered in black and white is called as a panchromatic image. A PAN image can be defined as a single band image generally displayed as shades of gray.

2.2.6.2 Multi-Spectral Images (MS)
Images acquired in more than one spectral or wavelength interval is defined as MS image. Each individual image is usually of the same physical area and scale but of a different spectral band. High-resolution, satellite-based imaging sensors provide one band of panchromatic data and four bands of multispectral data at a quarter of the resolution of the panchromatic data.

2.2.6.3 Pan Sharpened Multispectral Images (PS-MS)
This is not an actual image acquired from a satellite sensor. PS-MS image is an image generated by merging high resolution panchromatic and low resolution
multispectral imagery to create a single high resolution color image using any remote sensing software (fig. 2.6). For example, the SPOT satellite provides high resolution (1m) panchromatic data, while the LANDSAT satellite provides low resolution (30m) multispectral images. Image fusion attempts to merge these images and produce a single high resolution multispectral image.

For visual interpretation only three bands can be simultaneously displayed on computer monitors. Common band combinations are natural color red-green-blue (RGB) displays and Color Infra-Red (CIR) with Near Infra-Red (NIR) displayed as red, Red displayed as green, and green displayed as blue. Natural-color is good for analyzing man-made objects such as roads, buildings and bridges. Color-infrared combination, which is highly reflective in near IR, is often used for detection of vegetation and camouflage. Fig. 2.7 shows the same area in Panchromatic (Gray scale), Natural Color (RGB) and Color Infrared (CIR) combinations. We can see that in a CIR image, where infrared reflectance is displayed in red, the healthy vegetation appears bright red, while the rangeland remains quite low in reflectance.
2. Study of Feature Extraction Methods and a Survey of Existing Works

2.2.6.4 Radar Images

Images acquired by a radar sensor of a satellite are termed as radar images. Longer wavelength microwave radiation can penetrate through cloud cover, haze, dust, and all but the heaviest rainfall as the longer wavelengths are not susceptible to atmospheric scattering. This property allows a radar sensor to detect microwave energy under almost all weather and environmental conditions so that data can be collected at any time. Fig. 2.8 shows a cloudy area seen from an optical sensor and the same area seen from RISAT, the Radar Imaging Satellite launched by India.
2.2.7 Types of Satellites

Different satellites have different type of sensors which provide different type of images of the same region. Depending on the types of satellites there are mainly four categories:

- Meteorological satellites for weather monitoring and forecasting,
- Land observation Satellites for monitoring the Earth's surface,
- Radar imaging Satellites which enables imaging of the earth surface features during both day and night time and have cloud penetration capability, and
- Marine observation satellites for monitoring the Earth's oceans and water bodies.

Since the launch of Sputnik, thousands of satellites have been blasted off into space which falls into any of these categories. Of this, following subsections discuss about only a few number of satellites having application in this thesis.

2.2.7.1 Meteorological satellites

The images obtained from meteorological satellites are always multispectral images of low resolution. The most commonly used meteorological satellites for research purpose are MODIS — Moderate-resolution Imaging Spectroradiometer (http://modis.gsfc.nasa.gov), GOES — Geostationary Operational Environmental Satellite (http://goes.gsfc.nasa.gov) and AVHRR — Advanced Very High Resolution Radiometer (http://www.class.ngdc.noaa.gov) and the images can be directly downloaded. They were designed by National Aeronautics and Space Administration (NASA) for the National Oceanic and Atmospheric Administration (NOAA) to provide the United States National Weather Service with frequent, small-scale imaging of the Earth's surface and cloud cover. The GOES and AVHRR series of satellites have been used extensively by meteorologists for weather monitoring and forecasting for over 20 years. The different bands available for these sensors and their applications are briefed in table 2.1 & 2.2. MODIS is a payload launched into earth orbit by NASA in 1999 on board the Terra satellite, and in 2002 on board the Aqua satellite. The sensor capture data in 36 spectral bands ranging in wavelength from 0.4 µm to 14.4 µm and at varying spatial resolutions (2 bands at 250 m, 5 bands at 500 m and 29 bands at 1 km). Table 2.3 describes wavelengths represented by the 36 bands of MODIS sensor.
2. Study of Feature Extraction Methods and a Survey of Existing Works

Table 2.1: GOES Data

<table>
<thead>
<tr>
<th>Band</th>
<th>Wavelength Range (μm)</th>
<th>Spatial Resolution (m)</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.63-0.70 (visible)</td>
<td>1 km</td>
<td>Aerosol, pollution, sea surface reflectance, ocean storm identification</td>
</tr>
<tr>
<td>2</td>
<td>0.75-1.10 (near IR)</td>
<td>4 km</td>
<td>Identification of fog at night, discriminating water clouds and snow or ice clouds during diurnal cycles, detecting fires and volcanoes, night time assessment of sea surface temperature</td>
</tr>
<tr>
<td>3</td>
<td>0.67-0.72 (near IR)</td>
<td>4 km</td>
<td>Monitoring sea-ice extent, sea-ice thickness, and sea ice concentration</td>
</tr>
<tr>
<td>4</td>
<td>1.65-1.75 (middle IR)</td>
<td>4 km</td>
<td>Identification of low-level clouds, determination of sea surface temperature, detection of volcanic ash, and heavy snowfall</td>
</tr>
</tbody>
</table>

Table 2.2: AVHRR Data

<table>
<thead>
<tr>
<th>Band</th>
<th>Wavelength Range (μm)</th>
<th>Spatial Resolution (m)</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.68-0.69 (red)</td>
<td>1.1 km</td>
<td>Cloud, snow, and ice monitoring</td>
</tr>
<tr>
<td>2</td>
<td>0.725-1.11 (near IR)</td>
<td>1.1 km</td>
<td>Water, vegetation, and agriculture surveys</td>
</tr>
<tr>
<td>3</td>
<td>3.55-3.90 (mid IR)</td>
<td>1.1 km</td>
<td>Sea surface temperature, volcanoes, and forest fire activity</td>
</tr>
<tr>
<td>4</td>
<td>10.53-13.1 (thermal IR)</td>
<td>1.1 km</td>
<td>Sea surface temperature, soil moisture</td>
</tr>
<tr>
<td>5</td>
<td>11.5-12.5 (thermal IR)</td>
<td>1.1 km</td>
<td>Sea surface temperature, soil moisture</td>
</tr>
</tbody>
</table>

Table 2.3: MODIS Data

<table>
<thead>
<tr>
<th>Band</th>
<th>Wavelength (μm)</th>
<th>Resolution (m)</th>
<th>Primary Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.620-0.670</td>
<td>250</td>
<td>Land/Cloud/Atmosphere</td>
</tr>
<tr>
<td>2</td>
<td>0.841-0.866</td>
<td>250</td>
<td>Land/Cloud/Atmosphere</td>
</tr>
<tr>
<td>3</td>
<td>0.470-0.510</td>
<td>900</td>
<td>Land/Cloud/Atmosphere</td>
</tr>
<tr>
<td>4</td>
<td>0.545-0.655</td>
<td>500</td>
<td>Land/Cloud/Atmosphere</td>
</tr>
<tr>
<td>5</td>
<td>0.620-1.250</td>
<td>500</td>
<td>Land/Cloud/Atmosphere</td>
</tr>
<tr>
<td>6</td>
<td>0.628-1.672</td>
<td>500</td>
<td>Land/Cloud/Atmosphere</td>
</tr>
<tr>
<td>7</td>
<td>2.010-2.355</td>
<td>500</td>
<td>Land/Cloud/Atmosphere</td>
</tr>
<tr>
<td>8</td>
<td>4.05-0.420</td>
<td>1000</td>
<td>Ocean Color/Atmosphere</td>
</tr>
<tr>
<td>9</td>
<td>4.36-0.448</td>
<td>1000</td>
<td>Phytoplankton/Atmosphere</td>
</tr>
<tr>
<td>10</td>
<td>4.83-0.493</td>
<td>1000</td>
<td>Biogeochemistry/Atmosphere</td>
</tr>
<tr>
<td>11</td>
<td>5.26-0.536</td>
<td>1000</td>
<td>Ocean Color/Atmosphere</td>
</tr>
<tr>
<td>12</td>
<td>5.46-0.556</td>
<td>1000</td>
<td>Ocean Color/Atmosphere</td>
</tr>
<tr>
<td>13</td>
<td>6.62-0.672</td>
<td>1000</td>
<td>Ocean Color/Atmosphere</td>
</tr>
<tr>
<td>14</td>
<td>6.73-0.685</td>
<td>1000</td>
<td>Ocean Color/Atmosphere</td>
</tr>
<tr>
<td>15</td>
<td>7.43-0.735</td>
<td>1000</td>
<td>Ocean Color/Atmosphere</td>
</tr>
<tr>
<td>16</td>
<td>8.02-0.877</td>
<td>1000</td>
<td>Ocean Color/Atmosphere</td>
</tr>
<tr>
<td>17</td>
<td>8.90-0.920</td>
<td>1000</td>
<td>Ocean Color/Atmosphere</td>
</tr>
<tr>
<td>18</td>
<td>9.51-0.941</td>
<td>1000</td>
<td>Ocean Color/Atmosphere</td>
</tr>
<tr>
<td>19</td>
<td>9.15-0.965</td>
<td>1000</td>
<td>Ocean Color/Atmosphere</td>
</tr>
</tbody>
</table>

2.2.7.2 Land Satellites (LANDSAT)

Although many of the weather satellite systems are also used for monitoring the Earth's surface, they are not optimized for detailed mapping of the land surface. Driven by the exciting views from, and great success of the early meteorological satellites in the 1960's, the first satellite designed specifically to monitor the Earth's surface, Landsat-1, was launched by NASA in 1972. Landsat was designed as an experiment to test the feasibility of collecting multi-spectral Earth observation data from an unmanned satellite platform. Since that time, this highly successful program has collected an abundance of data from around the world from several Landsat satellites. Originally managed by NASA, responsibility for the Landsat program was
transferred to NOAA in 1983. In 1985, the program became commercialized, providing data to civilian and applications users. The most popular instruments of Landsat is the Multi Spectral Scanner (MSS) systems, and the Thematic Mapper (TM). Table 2.4 gives the MSS sensor details. The TM sensor provides several improvements over the MSS sensor including: higher spatial and radiometric resolution; finer spectral bands; seven as opposed to four spectral bands (USGS, 2001). The applications of TM bands are listed in table 2.5.

Table 2.4: Landsat MSS Data

<table>
<thead>
<tr>
<th>Channel</th>
<th>Wavelength Range (μm)</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSS 4, 5</td>
<td>0.5 - 0.6 (green)</td>
<td>soil/vegetation discrimination; coastal mapping; urban feature identification</td>
</tr>
<tr>
<td>MSS 5</td>
<td>0.6 - 0.7 (red)</td>
<td>green vegetation mapping (measures reflectance peak)</td>
</tr>
<tr>
<td>MSS 6</td>
<td>0.7 - 0.8 (near infrared)</td>
<td>vegetated vs. non-vegetated and plant species discrimination (plant chlorophyll absorption)</td>
</tr>
<tr>
<td>MSS 7</td>
<td>0.8 - 1.1 (near infrared)</td>
<td>identification of plant/vegetation types, health, and biomass content; water body delineation; soil moisture</td>
</tr>
</tbody>
</table>

Table 2.5: Landsat TM Data

<table>
<thead>
<tr>
<th>Channel</th>
<th>Wavelength Range (μm)</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM 1</td>
<td>0.45 - 0.52 (blue)</td>
<td>soil/vegetation discrimination; coastal mapping; urban feature identification</td>
</tr>
<tr>
<td>TM 2</td>
<td>0.52 - 0.60 (green)</td>
<td>green vegetation mapping (measures reflectance peak)</td>
</tr>
<tr>
<td>TM 3</td>
<td>0.63 - 0.69 (red)</td>
<td>vegetated vs. non-vegetated and plant species discrimination (plant chlorophyll absorption)</td>
</tr>
<tr>
<td>TM 4</td>
<td>0.76 - 0.90 (near IR)</td>
<td>identification of plant/vegetation types, health, and biomass content; water body delineation; soil moisture</td>
</tr>
<tr>
<td>TM 5</td>
<td>1.55 - 1.75 (short wave IR)</td>
<td>sensitive to moisture in soil and vegetation; discriminating snow and cloud-covered areas</td>
</tr>
<tr>
<td>TM 6</td>
<td>10.4 - 12.5 (thermal IR)</td>
<td>vegetation stress and soil moisture discrimination related to thermal radiation</td>
</tr>
<tr>
<td>TM 7</td>
<td>2.08 - 2.35 (short wave IR)</td>
<td>discrimination of mineral and rock types; sensitive to vegetation moisture content</td>
</tr>
</tbody>
</table>

2.2.7.3 High Resolution Satellite Imaging Sensors

The images used in this study of structural feature extraction are obtained from CARTOSAT (India), IRS-1D (India), Quickbird (USA), IKONOS (USA) and SPOT.
(France) sensors because of its high resolution and the available multispectral bands which are also of good resolution compared to the Landsat series. Table 2.6 gives the details of these sensors.

Table 2.6: High Resolution Imaging Sensors

<table>
<thead>
<tr>
<th>SATELLITE</th>
<th>SENSORS</th>
<th>SPATIAL RESOLUTION (m)</th>
<th>SPECTRAL RESOLUTION (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRS (Indian Remote Sensing Satellite)</td>
<td>IRS – 1D</td>
<td>Panchromatic: 5.8</td>
<td>Panchromatic: 0.50 - 0.75 Multispectral: Green: 0.52 – 0.59 Red: 0.62- 0.68 Near-IR: 0.77- 0.86 Short Wave -IR: 0.77- 0.86</td>
</tr>
<tr>
<td>CARTOSAT (Cartographic Satellite)</td>
<td>Cartosat- 2</td>
<td>Panchromatic: 1</td>
<td>Panchromatic: 0.50 - 0.75</td>
</tr>
<tr>
<td>QuickBird</td>
<td>QuickBird</td>
<td>Panchromatic: 0.61 Multispectral: 2.44</td>
<td>Panchromatic: 0.445 -0.90 Multispectral: Blue: 0.45 – 0.52 Green: 0.52 – 0.60 Red: 0.63- 0.69 Near-IR: 0.76- 0.90</td>
</tr>
<tr>
<td>IKONOS</td>
<td>IKONOS</td>
<td>Panchromatic: 1 Multispectral: 4</td>
<td>Panchromatic: 0.45 - 0.90 Multispectral: Blue: 0.45 - 0.53 Green: 0.52 - 0.61 Red: 0.64 - 0.72 Near-IR: 0.77 - 0.88</td>
</tr>
<tr>
<td>SPOT 1, 2, 3 (Satellite Pour l'Observation de la Terre)</td>
<td>SPOT 1 SPOT 2 SPOT 3</td>
<td>Panchromatic: 10 Multispectral: 20</td>
<td>Panchromatic: 0.50 - 0.73 Green: 0.50 - 0.59 Red : 0.61 - 0.68 Near-IR : 0.78 - 0.89</td>
</tr>
<tr>
<td>SPOT 4, 5</td>
<td>SPOT 4 SPOT 5</td>
<td>SPOT 4: Panchromatic: 10 Multispectral: 20</td>
<td>SPOT 4: Panchromatic: 0.61 - 0.68 Green : 0.50 - 0.59 Red : 0.61 - 0.68 Near-IR : 0.78 - 0.89 Short-wave IR (SWIR) : 1.58 - 1.75</td>
</tr>
</tbody>
</table>
2.2.8 Satellite Image Processing Softwares

We will never get a satellite image as such; instead it is available as digital data having panchromatic or multispectral information. Different bands have different types of information for processing. Also depending on different levels of information, the processing complexity varies. Though MATLAB has an image processing toolbox, some other specialized softwares are required for the processing of multi-band remotely sensed data at different levels. Some of the softwares are discussed here:

2.2.8.1 ENVI (Environment for Visualizing Images)

ENVI (http://www.exelisvis.com/ProductsServices/ENVI/ENVI.aspx) is a software application used to process and analyze geospatial imagery. It combines file-based and band-based techniques with interactive functions. Some of these functions include data transforms, filtering, classification, registration and geometric corrections, spectral analysis tools, and radar tools. ENVI does not impose limitations on the number of spectral bands that can be processed, so multi-spectral or hyperspectral data sets can be used. ENVI also includes many advanced functions that allow for analysis of radar data sets (ENVI, 2004).

2.2.8.2 Erdas Imagine

ERDAS Imagine (http://geospatial.intergraph.com/products/ERDASIMAGINE/ERDASIMAGINE/Details.aspx) is aimed primarily at geospatial raster data processing and allows the user to prepare, display and enhance digital images for mapping use in GIS. By manipulating imagery data values and positions, it is
possible to see features that would not normally be visible and to locate geo-
positions of features that would otherwise be graphical. The level of brightness or
reflectance of light from the surfaces in the image can be helpful with vegetation
analysis and other feature extraction applications.

### 2.2.8.3 PCI Geomatica

PCI Geomatica ([http://www.pcigeomatics.com/software/geomatica2013](http://www.pcigeomatics.com/software/geomatica2013)) is a remote
sensing software package for processing earth observation data, designed by PCI
Geomatics Inc. This software is aimed primarily at raster data processing and allows
users to load satellite and aerial imagery where advanced analysis can be performed.

### 2.2.8.4 MultiSpec

MultiSpec ([http://cobweb.ecn.purdue.edu/~biehl/MultiSpec](http://cobweb.ecn.purdue.edu/~biehl/MultiSpec)) is a freeware
 multispectral image data analysis system. Multispec can be used to analyze
multispectral and especially hyper spectral image data such as LANDSAT, MODIS,
AVHRR etc. As there are many unsupervised clustering algorithms available in this
software, it is best suitable for image interpretation.

### 2.2.9 Image Interpretation using Landsat Data

Eventhough the availability of high resolution sensors, Landsat Thematic Mapper is
still the best one for image interpretation. The features for interpretation can be
‘water bodies’, ‘soil boundaries’, ‘urban areas’, ‘agricultural areas’ or any other.
Degree of identification of the feature using a band can be ‘good’, ‘medium’, ‘poor’,
or ‘the particular band is ‘not suitable’. We can also use band combinations for
identification of features. Using the combinations we can arrive at a table of
‘features’ and a similar degree of identification or a ranking pattern. The following
table give such a guide for the Landsat TM sensor taken from FAO (1988).

<table>
<thead>
<tr>
<th>Themes/TM Bands</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water bodies</td>
<td>G</td>
<td>P</td>
<td>M</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Water characteristics</td>
<td>G</td>
<td>G</td>
<td>P</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Drainage patterns</td>
<td>P</td>
<td>P</td>
<td>M</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Soil boundaries</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>G</td>
<td>G</td>
<td>M</td>
</tr>
<tr>
<td>Forested area</td>
<td>P</td>
<td>M</td>
<td>M</td>
<td>G</td>
<td>G</td>
<td>M</td>
</tr>
<tr>
<td>Agricultural areas</td>
<td>P</td>
<td>M</td>
<td>M</td>
<td>G</td>
<td>G</td>
<td>M/G</td>
</tr>
<tr>
<td>Urban area</td>
<td>M/G</td>
<td>G</td>
<td>G</td>
<td>P</td>
<td>P</td>
<td>PM</td>
</tr>
<tr>
<td>Quarries</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
</tbody>
</table>

G: Good
P: Poor
N: Not usable
M: Medium
Though the table given above will give some approximate assumptions of the type of region under study, it is necessary to find the surface temperature of the land cover to estimate the type of region (Van, T.T., 2005). Accordingly the relationship between surface temperature and land cover types is obtained as shown in fig.2.9.

![Figure 2.9: Thermal signature of different regions](image)

From the figure it is understood that a lot of building is one of the factors that more heat reflection occurs and it will raise the surface temperature at urban area. The cooler areas with temperature between 24°C to 28°C are areas with abundant vegetation (grow-up paddy field, forest). This is the result of dissipating solar energy by absorbing surrounding heat and through an evaporation process from the leaves. Thus, finding the relationship between the surface temperature and land cover type will give an idea of different regions and is able to mask the unwanted regions so that the region of interest will be projected and the feature extraction process will be easier.

### 2.2.10 Digital Elevation Model (DEM)

Digital Elevation or Digital Terrain Model (DEM/DTM) is the digital representation of elevation data of the ground surface. Or we can say, DEM is contour lines of equal elevation. DEM of the region will give the height information of that region. Generating DEM from Aster Level L1A data and Cartosat-1 images are explained in Kayadibi, O. (2009). SPOT-5 sensor has a high resolution stereoscopic (HRS) instrument to facilitate preparation of digital elevation models (DEMs) at a resolution of 10m. The method can be applied on SPOT-5 data for better results. This height information obtained from DEM is another supporting tool for image
interpretation. The result obtained in Kayadibi, O. (2009) is shown below (fig. 2.10) to see how the information about a lake is obtained from DEM of the region. The elevation of lake will be high when compared to other regions.

![Figure 2.10: (a) Generated DEM (b) 3D Viewing of the DEM [courtesy: Kayadibi, O. (2009)]](image)

### 2.2.11 Identifying Vegetation and Water areas using High Resolution Sensors

Masking off the vegetation and water areas is very useful for the case of structural feature extraction because if a region is identified as vegetation/water area, it will never possible to be a structural feature area. Two very common indices are there in remote sensing to attain this task – Normalized Differential Vegetation Index (NDVI) and Normalized Differential Water Index (NDWI).

Examining the ratio of reflected infrared to red wavelengths is an excellent measure of vegetation health. This is the premise behind NDVI which is derived by Rouse, J. W et al. (1973). Vegetation areas have a high NDVI value because of their high reflectance of infrared light, and relatively low reflectance of red light. Computation of NDVI is shown in eqn. 2.1.

\[
NDVI = \frac{NIR - R}{NIR + R} \tag{2.1}
\]

The design of a spectral water index was based on the fact that water absorbs energy at near-infrared (NIR) quite opposite to vegetation areas. Therefore, the NIR reflection values of water regions are generally smaller than the other regions. This concept is used for calculating NDWI by McFeeters, S.K. (1996). Computation of NDWI is shown in eqn. 2.2.
\[ NDWI = \frac{G - NIR}{G + NIR} \]  

(2.2)

Since the water regions have higher reflection values in green band than near infrared band, NDWI values of water regions will be higher than the rest.

### 2.3 Feature Extraction

In computer vision or image understanding, the concept of feature extraction refers to methods that aim at computing abstractions of image information and making local decisions at every image point, whether there is an image feature of a given type at that point or not and extracting it. A feature is defined as an ‘interesting’ part of an image, and features are used as a starting point for all computer vision algorithms. Since features are the main primitives for subsequent image understanding algorithms, the overall algorithm will often only be as good as its feature extraction system. According to Haralick, R. M. (1992), image understanding consists of several steps: First of all comes image acquisition. The second step is image pre-processing that aims at image enhancement through noise reduction, contrast enhancement, sharpening, geometric correction, radiometric calibration etc. The last step - image analysis or image interpretation - leads to a scene description after extracting relevant features from the image. Image analysis itself consists of several subtasks:

- segmentation (labeling, grouping) to identify image objects (i.e. segments, edges, lines, pointlike structures) that have a meaning in the real word;

- extraction to provide the identified segments with a list of properties (spectral, spatial, textural properties, etc.);

- classification (or matching) to give a meaning to the segments, i.e. to relate real world objects to image objects (i.e. segments).

Many image understanding methods incorporate these steps explicitly. In addition, depending on the application, the above stated sequence of steps may have to be applied at one or more levels of the recognition process. Since this work mainly deals with feature extraction, this topic is singled out and reviewed in more detail.
2. Study of Feature Extraction Methods and a Survey of Existing Works

2.3.1 Techniques for Feature Extraction

There is a wide range of techniques used to detect features in imagery. To carry out object recognition, it is first necessary to establish a model or framework that describes the general characteristics of the feature of interest (Suetens, et al., 1992). Automated feature extraction requires that such a model be defined in a manner that can be implemented by computer (Trinder & Wang, 1998). Model based processing exploits the constraints and relationships that define objects, for example, the size, shape and material of a building, or the width, material, and direction of a road/bridge. The feature model includes information relating to a range of characteristics such as intensity, shape, texture, and context. An objective function is used to find a best fit between the model and the image data. Some techniques use hierarchical types of models (Suetens, et al., 1992). For example, Yee, B. (1987) identified bridges by first finding potential road segments, then restricting the search to select those with water on either side.

The simplest models rely only on local intensity values to recognize features, as is the case in a traditional supervised classification. Suetens et al. (1992) believe that without considering geometric and semantic characteristics as well as statistical properties, most feature extraction procedures are unlikely to succeed. Baumgartner et al. (1999) initially defined three contexts in order to subdivide a scene: urban, rural, and forest. Road sub-models were then developed locally within each of the global contexts, reflecting the complex nature of a typical feature. Katartzis et al. (2001) used a model that combined both geometric and radiometric properties of the features they aimed to extract. Developing models of features to be extracted can provide significant processing benefits.

Though there are different methods and models available for structural feature extraction like roads, buildings and bridges, it is desirable to mask shadows and clouds from the region before going for the feature extraction process. Shadows and clouds may cause many undesirable problems such as loss of feature information, false color tone and shape distortion of objects, which seriously affect the quality of images, and directly influences computer vision and image analysis tasks, such as edge detection, image segmentation, object recognition, video surveillance, and stereo registration. The amount of shadows and clouds increases with the spatial resolution and so masking shadow/cloud infected pixels is important.
before going for any feature extraction processes for the case of high resolution satellite images.

The level of automation in the reviewed techniques varied significantly. Some procedures require a significant amount of human input to select potential feature locations. Other procedures use a few initial assumptions, such as the relative brightness of the pixels to be extracted, and allow the computer to do the rest. A common application of structural feature extraction is updating GIS data layers. As a result, increasing numbers of techniques are taking advantage of this by using existing digital data to provide preliminary input information for processing.

Many of the techniques reported in the literature combine strategies from a variety of approaches. Categorizing such approaches becomes a challenge. After reviewing about shadow and cloud masking, different strategies used for the structural feature extraction like roads, buildings and bridges are elaborately discussed in this chapter.

2.3.1.1 Methods for Shadow Masking

Recently the advent of optical sensors with very high resolution can supply more details of the earth surface condition. However, the effects of shadow in these images are remarkable. Since shadow changes depending on the time and season, it would be extracted as errors in change detection of the earth surface. This is particularly acute in cities where dense high-rise buildings cast many long shadows across a variety of different surface types. Due to the darker shadow pixels, feature identification becomes a problem when the same surface having large color differences. The negative effect of shadows including radiometric information loss (Arevalo et al., 2006; Chen et al., 2007) and inaccurate change detection results due to different azimuth angle (Yamazaki et al., 2009). In order to minimize the negative effect of shadows in aerial photos and satellite images, the concept of brightness compensation was introduced (Shu & Freeman, 1990). This process requires shadows firstly be detected and extracted, and then different algorithms applied to enhance these shadow regions.

Shadow detection techniques can be categorized as properties based and model based. For properties based detection, Cai et al. (2010) suggested 3 different indices based on HSI (Hue Saturation Intensity) color space and NDVI index for
2. Study of Feature Extraction Methods and a Survey of Existing Works

shadow detection from high resolution remote sensing images which is more accurate than using HSI color space alone. Tsai (2006) compared the effectiveness of different invariant color spaces for shadow detection, including HSI, HSV (Hue Saturation Value), YIQ (Luminous, In-phase & Quadrature components) and YCbCr (Luminous, Blue-difference & Red-difference Chroma components) models. He concluded that HSI, YIQ and YCbCr are the most suitable color space for shadow detection. However, when these algorithms were applied to dense urban area, shadows are usually mixed with water bodies which may get wrongly enhanced. In Chen et al. (2007), 5 empirical indices were suggested to separate shadows from water and is named as Spectral Shape Index.

Model based shadow detection algorithms require a Digital Building Model (DBM) together with Digital Surface Model (DSM) (Zhou, G., 2005). Nakajima et al. (2002) made use of Airborne Laser Scanning data and DSM for shadow simulation in urban area. However, a high accuracy DSM is required to produce fair result. In Hong Kong, an accurate DSM can hardly be obtained due to the fact that high-rise buildings are densely distributed on mountainous terrain. In the method proposed by Liu & Yamazaki (2010) the characteristics of radiance ratio (shadow/sunlit) in Quickbird images are investigated by radiance measurements and the effective spectral quantum efficiency of the sensor. Then shadowed areas are extracted from a Quickbird image by an object-based method and the radiance ratio of shadow and sunlit areas. Later they have proposed another algorithm based on radiance measurement (Liu & Yamazaki, 2012). However, the distinction of colors between the shadows of buildings and water surface is still a challenge to the professionals (Nath, R.K. & Deb, S.K., 2009) which is investigated in our work and tried to eliminate this confusion quite effectively.

2.3.1.2 Methods for Cloud Masking

During the last decade many approaches have been proposed to retrieve satellite images from clouds and cloud shadows. Dell’Acqua, F. & Gamba, P. (2001) presented a tool for shape similarity evaluation for query-by shape searching into meteorological image archives based on the point diffusion technique. Holowczak, R. et al. (2002) reported a system that can automatically determine whether a region of interest is visible in the image, free from cloud, and can incorporate this into the meta-data for individual images to enhance searching
Nauss, T. et al. (2005) have proposed an algorithm based on the analytical solutions of the radiative transfer equations valid for optically thick weakly absorbing cloud layers. Fu, D. & Xu, L. (2011) have used 2D-Gabor wavelet in satellite image classification having clouds. Upreti, D. (2011) has used Gray level Co-occurrence Matrix (GLCM) and histogram quantization technique to retrieve cloud patterns to discover tropical cyclone. Most of the previous work was directed to meteorological observation images with very low resolution which is not applicable to high resolution satellite images. Also it doesn’t care with cloud removal preprocessing operation which is still done manually.

The remaining sections of this chapter present different techniques used in automatically extracting structural features from satellite imagery, recognizing that there is often substantial overlap between the procedures.

2.3.1.3 Edge detection

Edge detectors seek to find areas in images where the brightness changes rapidly over a short distance. The points where these changes occur are marked as edge points. A common edge detection approach begins by detecting areas within a local window where a considerable change in intensity occurs. The change in intensity at a given point might be gradual and span over a few pixels. A thinning algorithm can be used to reduce this area to a single point. The end result is a line running along the edge where the change in intensity is the greatest. This simplistic edge detection approach produces localized edges and is sensitive to noise or disruption along the edge line. In general, road and bridge objects appear as linear structures at lower spatial resolutions and as homogeneous elongated segments in higher spatial resolution (Wiedemann, C. & Hinz, S., 1999; Bai, Z. et al., 2005). Most of the buildings are rectangular in shape and can be considered as a four dimensional linear structure (Wang, Y. et al., 2008). Edge detectors can be applied to lower resolutions to extract linear features or to elongated segments in higher resolutions which supports the structural feature extraction process. The versatility of these algorithms has made them very popular in feature extraction literature both in spatial domain and in spectral domain (Patterson, T. J., 2002). In general, edge detectors have been used to provide information to higher level processes such as segmentation, contextual and geometrical classifiers. Edge detection is a very popular tool, but is
not a definitive solution to automatic feature extraction since it is sensitive to noise and occlusions.

### 2.3.1.4 Hough transform

The Hough Transform (HT) (Hough, P.V.C, 1962) is a technique used to find features of a particular shape in digital imagery. The classical HT is used to detect regular shapes, such as lines, circles and ellipsoids. The generalized HT is a parametric approach where a line segment is defined in its parametric notation (eqn. 2.3) as first introduced by Duda, R. & Hart, P. (1972).

\[ x \cos \theta + y \sin \theta = r \]  

(2.3)

where \( r \) denotes the length of a normal drawn from the origin to the line of interest and \( \theta \) is the orientation of the line measured from the x-axis.

Therefore the image from the (x,y) domain is transformed into (r, \( \theta \)) domain. Within this domain one then looks for significant peaks which correspond to strong linear features of the image. Selecting these peaks the linear features can be reconstructed in the (x,y) domain through an inverse transformation. During this transformation pixels where linear features are identified are assigned one and the others are assigned a zero value.

One of the advantages of using the HT is that it is relatively unaffected by noise and fragmentation. Geman, D. & Jedynak, B. (1991) are among the first to use the HT within a road network extraction system. Applicability of hough transform for region extraction in IRS images is discussed elaborately by Pal, S.K. et al. (1998). Jin & Davis (2005) developed the Spatial Signature Weighted Hough Transform (SSWHT) to detect the grid-like shape of roads in dense urban scenes.
The SSWHT uses the length of the spatial signature in a given scene as a weighting for the HT rather than simply using a value of one as weighting. Because bridges in images are always straight lines, the classical Hough transformation is employed for extracting straight lines from the candidate bridges image. Utilizing this concept most of the bridge detection algorithms uses the advantage of HT in detecting lines at any of the stage. Hao, Qiwei. *et al.* (2007) proposed a fast HT for automatic extraction of bridges exclusively for visible light satellite images with more texture and complex scenes. Sui, H. *et al.* (2006) used the same HT for the NIR band of multispectral IKONOS images to extract bridge pixels.

As most of the buildings are rectangular, the method using HT can be used for the detection of buildings also. Tao, W. B. *et al.* (2002) suggested a windowed HT in which four extracted peaks satisfy certain geometric conditions are considered as building. The method is extended to any type of buildings irrespective of their shape in recent researches using modified HT (Wang, C.K. *et al.*, 2007; Pakizeh, E. *et al.*, 2010). Koc San, D. & Turker, M. (2010) used rectangular and circular HTs to detect all rectangular and circular shaped buildings and most of the buildings having shapes in between these.

**2.3.1.5 Mathematical Morphology**

The techniques of mathematical morphology have proven useful in automating feature extraction. Destival (1986) and O’Brien (1989) used mathematical morphology to search for roads in simulated and actual SPOT imagery, respectively. In mathematical morphology, images are filtered using a kernel. The output of the filtering process depends on the match between the image and the kernel and the operation being performed. The two basic operations of mathematical morphology are dilation and erosion (Serra, J., 1986).

Dilation followed by erosion (closing) closes small gaps and connects sets; erosion followed by dilation (opening) removes small or narrow elements, without effecting large ones (Dong, P., 1997). Dilating the gray scale image fills holes and trenches, broadens peaks and ridges, and shifts steps outward. Erosion tends to eliminate features with a width less than the kernel. Ansoult *et al.* (1990) and Yamada *et al.* (1993) used mathematical morphology as a means of acquiring GIS layers from scanned maps. In order to find closed boundaries of the various classes
displayed on the maps, Ansoult et al. (1990) used a mathematical morphology routine called a watershed algorithm. This algorithm is performed in two stages, a skeletonization followed by a pruning. The skeleton is formed by sequentially thinning the image with structuring elements that preserve homotopy; the pruning transformation removes lines that do not form closed contours. These two operators are most easily explained by referring to the illustration shown in fig.2.12.

Géraud & Mouret (2004) applied a segmentation algorithm to grayscale satellite imagery. The images contained numerous local minima which resulted in over-segmentation. The mathematical morphology closing operation is ideally suited to suppress local minima.

![Figure 2.12: Stages of Watershed Algorithm (a) Initial (b) Skeletonized (c) Pruned](image)

The disc shape has the drawback of creating artefacts when the filtering strength increases, for example, peaks tend to shift when a large number of minima are removed. Vincent, L. (1992) used an area closing operation which does not have the same deficiencies as the classical closing operation. Chanussot, J. et al. (1999) found that morphological operators perform well for extracting roads because they are intrinsically characterized by shape. Roads are identified using mathematical morphology by considering the intensity of pixels in relation to their neighbors. The procedure uses a priori knowledge about the relative intensity of roads to perform the automated road extraction, for example, assuming that roads have brighter intensity than the background (Trinder & Wang, 1998). Later many works for road network extraction used morphological operations in the segmentation phase effectively. Gaetano, R. (2011) developed a morphological road segmentation algorithm particularly for urban areas. The method uses the possibilities of skeletonization and pruning, watershed transform, morphological opening and
closing for efficient segmentation. A semi-automated road network extraction method using directional morphological enhancement and segmentation techniques by Chaudhuri, D., et al. (2012) is evaluated on a variety of images from IKONOS, QuickBird, CARTOSAT-2A satellites to validate the accuracy and efficiency of morphological segmentation in feature extraction. Though bridges have many similar features like roads, works based on morphological operations in bridge detection is not much discussed. The works done by Kartal, M. (2004), He & Liu (2010) and Gedik, E. et al. (2012) are some of the techniques presented in literature in which segmentation based on mathematical morphology is part of the algorithm.

Building extraction is one of the biggest problems in the field of photogrammetry and remote sensing because of the irregular shapes and inconsistent features to be extracted. But this inhomogeneity leads to the importance of morphological operations with different structuring elements of varying sizes and shapes. A number of methods are proposed by researchers working in the field of remote sensing using the concept of morphology for building extraction. Benediktsson et al. (2005) first proposed this method using mathematical morphological operations to extract structural information from the image. Features generated by a Differential Morphological Profile (DMP) were selected by discriminant analysis and decision boundary feature extraction. Buildings and other land use categories were then classified using a neural network. Later Jin & Davis (2005) presented an automated building extraction strategy using the same DMP profile but simultaneously exploited structural, contextual, and spectral information from the image. After that many algorithms were developed with certain modifications (Lefevre, S., 2007; Aytekin, O., 2009; Xue, L., 2012). Huang, X. & Zhang, L. (2011) presented an efficient and accurate multidirectional and multiscale Morphological Index for automatic building extraction from multispectral satellite images where the overall accuracy reaches a percentage of 93.2.

2.3.1.6 Multi-Resolution Analysis (MRA)

According to Gonzalez & Woods (2008), MRA theory deals with the representation of images (or signals) in different resolutions (or frequencies). MRA is useful when applied to feature extraction processes in satellite images since certain features become more apparent at different resolutions. Fig. 2.13 illustrates what the Canny edges for the corresponding images in the image pyramid would look like. It is clear
that certain features are more apparent at different levels. It would be feasible to apply different extraction techniques at different resolutions, with the aim of increasing the accuracy of feature extraction systems. MRA is also beneficial in reducing the processing time of feature extraction systems. A common approach begins by detecting regions of interest at a coarse level and moving down to a finer level only in these areas.

A great variety of methods are found in literature whereby remotely sensed imagery is translated into a number of given sub-levels with the aim of increasing the accuracy of feature extraction systems. One of the earliest MRA road extraction approaches is presented by Steger et al. (1995) where a different extraction method is used for each scale level. More examples of the use of scale independent road extraction algorithms can be found in Lee, H. Y. et al. (2000); Dell’Acqua, F. et al. (2002); Heller J. & Pakzad, K. (2005).

Another popular MRA approach is based on the well known wavelet transform. Wavelet transforms can be applied to feature extraction in two different ways (Zhang, Q. & Couloigner, I., 2004). The first is to derive multiple image resolutions through a wavelet transform. Such an example is presented in the work of Gruen & Li (1995), where the wavelet acts as an impulse response of a band-pass
filter. The second approach is to perform analysis on the wavelet domain itself. 

Couloigner & Ranchin (2000) present an algorithm where an edge detector is applied in the wavelet domain. Many works are still developing using multi resolution analysis especially using wavelet filters in the field of feature extraction. Naouai, M. et al. (2011) proposed a new approach for road network extraction based on a quaternionic wavelet transform. Yong, B. et al. (2009) combined MRA with snake model (Kass, M. et al., 1988) for increasing the efficiency. Combining wavelet transform with watershed segmentation is another efficient method for feature extraction from satellite images (Parvathi, K. et al., 2009). Saroj, K. et al. (2007) inferred that utilizing the extracted features obtained by the wavelet transform (WT) rather than the original multispectral features of remote-sensing images provide better results. Due to the diversity of algorithms which is possible to develop using MRA, the method is extended to building and bridge extraction also. In building extraction, an investigation into the use of wavelets and scaling for the extraction of buildings in aerial images is first discussed by Levitt, S. & Aghdasi, F. (1998). Later the method is extended to satellite images for the detection of buildings (Tokunaga, M. & Tuong Thuy Vu, 2007; Liu, Z., 2008). Recent works in bridge detection using MRA and wavelet transforms are the works done by He & Liu (2010) and GurayErus et al. (2010). GurayErus et al. (2010) suggested a hybrid structural/statistical approach which aims to construct an intermediate step between the low-level image characteristics and high-level semantic concepts and recognizes structured and compact objects like bridges or roundabouts.

2.3.1.7 Classification

A classifier is a type of an inference model that seeks to identify patterns in data. Wasserman, P. (1983). Classifiers are typically used to categorize data into different pattern classes through labelling or by assigning probability values to data samples. These probability values indicate the likelihood of an observation belonging to a pattern class. Classifiers can also be used to extract features from Remote Sensed (RS) imagery and are often used to identify, amongst others, vegetation types, land usage, man-made objects, and natural features. Classification can be formalized as the process of applying decision functions to a set of unknown patterns. Consider the set \( x = (x_1, x_2, ..., x_n)^T \) with \( n \) patterns and \( W \) pattern classes \( \omega_1, \omega_2, \ldots, \omega_W \). The
objective of the classifier is to find \( W \) decision functions \( d_1(x), d_2(x), \ldots, d_W(x) \) that, if a pattern \( x \) belongs to class \( \omega_i \), then,

\[
d_i(x) > d_j(x) \quad j = 1, 2, \ldots, w ; j \neq i
\]  

(2.4)

The unknown pattern \( x \) would be labelled as being part of pattern class \( i \), if \( d_i(x) \) yields the largest value of all decision functions. From this definition, it is evident that the classification accuracy is directly correlated to the ability of the decision function to represent the pattern classes. Decision functions can model the pattern classes more accurately when the classes are separable. Classifiers can be used to categorize any type of data, but are often used in feature extraction system to perform any of the following:

- Spectral classification
- Textural classification
- Geometrical classification
- Contextual classification

2.3.1.7.1 Spectral classification

Spectral classifiers are used to compare the spectral signature of a pattern class to that of an unknown sample. A major issue with classifiers relates to the overlap that occurs between features in RS imagery. Considering that various features might have the same spectral signals, the classifier would be unable to make a clear distinction. This limitation can be reduced by considering additional information such as texture, structure and context.

A number of different classification methods are used in literature and include, amongst others, Artificial Neural Networks (ANN) and Fuzzy classifiers. The following sections discuss the manner in which these and other methods are employed for spectral classification.

2.3.1.7.1.1 Artificial Neural Network Classifiers

An ANN is a non-linear statistical model, which is based on biological neural networks (Wasserman, P., 1983). The ANN model consists of a multitude of interconnected non-linear components, which are known as neurons. In general, ANNs are dynamic systems able to ‘learn’ by adapting their internal structures according to information that flows through the network. Road network extraction
using ANNs is discussed even before the availability of high resolution images. The advantages of neural networks in the field of remote sensing is first discussed by Atkinson, P.D. & Tatnall, R. L. (1997). In the year 2000, Xiangyun, Hu. et al. had done the first work which is a semi automated approach of road extraction from aerial images. Mokhtarzade & Zoje (2007, 2008) thoroughly investigated ANN road extraction system and proposed different methods that classify Very High Resolution (VHR) imagery such as that obtained by the Quickbird and IKONOS satellites. The study also considers several different spectral ANN classifier configurations with the intent of finding the optimal road pixel classification solution. Each configuration is presented with a training set of pattern classes and is trained with the well-known steepest gradient descent algorithm Bishop, C. M. (1995). The ANN is tested with different input vectors and varying hidden layer sizes. All pattern classes are spectral and created from the RGB bands. Each structure has one output indicating whether the considered pixel is a road pixel or not. Recent works using this approach is found in the works of Mangala, R. (2010), Benkouider, F. et al. (2011) and Ghasemloo, Nima. et al. (2013).

Priestnall, G. et al. (2000) extracted urban features from 2m resolution Light Detection & Ranging (LiDAR) data in which ANN is used to discriminate between buildings and trees. The advent of LiDAR data has opened a new phase of building detection and city surface modeling research. LiDAR provides point clouds that significantly improve the accuracy of building detection (Alharthy, A. & Bethel, J., 2002) and highlight the importance of surface information on the building modeling and extraction process. Lari, Z. & Ebadi, H. (2007) had proved that combining spectral and structural information with ANNs provide very good efficiency for the case of building extraction from satellite images. A different approach using pulse coupled ANNs in very high resolution SAR images for building extraction is discussed by Frate & Pratola (2009) but applicable only to low density urban areas. The very common method for bridge detection using ANNs rely on radiometric features and neural networks to classify each pixel into several terrain types, and then using fixed rules to find bridges in this classification (Trias-Sanz, Roger., 2003). In the work of Zhang, Jingfei. (2011), after region of interest segmentation hopfield neural networks is utilized to recognize bridges in the final stage.
2. Study of Feature Extraction Methods and a Survey of Existing Works

2.3.1.7.1.2 Fuzzy classifiers

Fuzzy classifiers are considered to be classifiers based on fuzzy logic or fuzzy sets (Zadeh, L. A., 1965), which are based on set theory where the elements have a varying degree of membership. Fuzzy logic is expressed in terms of two truth values (true and false) while fuzzy sets allow multi-valued logic. As with classical ANN classifiers, fuzzy classifiers derive information for their membership functions from predetermined pattern classes and use the information to classify unknown patterns accordingly. Melgani, F. et al. (2000) proposed a method for fuzzy classification in multispectral satellite images. Mohammadzadeh et al. identify five pattern classes and define a representative spectral membership function for each class. The membership functions are constructed by calculating the mean and standard deviation values in each band of every pattern class.

Shackleford & Davis (2003) also use spectral fuzzy classifiers, but they detected a significant overlap between the spectral values of the pattern classes (tree and grass, road and building). To discriminate between these classes, a hierarchical fuzzy classifier is constructed. The hierarchical structure allows additional information, specific to each class, to be used. Textural information is used to distinguish between the tree and grass classes, and geometrical information for the road and building classes. Tuncer, O. (2007) follows the same approach and obtained better results by adding geometrical information to the spectrally derived membership functions. Levitt, S. & Aghdasi, F. (2000) applied fuzzy classification in building detection from remotely sensed images. Sumer, E. & Turker, M. (2008) used an adaptive fuzzy-genetic approach for building detection which is very accurate at that time but much time consuming. Yousefi, Bardia. et al. (2012) segmented building roofs according to the fuzzy membership functions, obtained from structural context. Erener, A. (2013), combined fuzzy classification with shadow influence for the efficient detection of buildings.

2.3.1.7.2 Textural classification

Textural classifiers endeavour to classify image regions according to their textural characteristics. To classify these regions, texture has to be mathematically quantified. Even though no formal definition for texture exists, descriptors such as smoothness, coarseness and regularity are often used (Hall, D. & Ball, G., 1965). One of the most celebrated works on texture analysis was written by
Haralick, R.M. (1979) in which 14 textural features are defined. A later study by Ohanian & Dubes (1992) reveals that three of these features are sufficient for texture classification, namely angular second-moment, contrast and entropy. As with most of the aforementioned spectral classifiers, texture classification begins by firstly defining pattern classes, which are then used to classify unknown patterns. The approach, proposed by Mao & Jain (1992), used a texture cube to calculate the similarity between pairs of pixels. A texture cube is created by considering the $3 \times 3$ neighbourhood around a candidate pixel (unknown pattern). This neighbourhood constitutes the width and height of the cube, while the depth consists of the RGB bands. Co-occurrence matrices are constructed to interweave the spectral and textural information contained within the cube. The following Haralick features are used as texture descriptors: correlation, energy, entropy, maximum probability, contrast, and inverse difference moment. A distribution from the six features is created for both the pattern class and the unknown pattern (candidate pixel). The distance between the two distributions is computed with the Bhattacharyya distance function (Bhattacharyya, A., 1945). The resulting distance is transformed to represent a pseudo-probability of the likelihood of a pixel being a feature pixel.

Haralick features are also used by Cai et al. (2005) to train a decision tree and a Support Vector Machine (SVM) classifier to classify road features. Zhang, Yun. (1999) suggested texture filtering for building detection which is one of the best works in this area. Dial et al. (2001) create a panchromatic texture filter that calculates the variability of the area around a sample pixel through an Angular Texture Signature (ATS). The information gathered by ATS is then passed to a higher level process for further classification. Textural classification is a popular approach for detecting manmade objects from satellite images and is used in studies by Samadzadegan, F. et al. (2008), Zhang, Q. et al. (2006), Cetin, M. et al. (2010), Miura, Hiroyuki. et al. (2012) and Benarchid, O. et al. (2013) amongst others.

2.3.1.7.3 Geometrical classification
Geometrical classifiers detect features based on distinctive structural characteristics. Edge detectors, segmentation algorithms or the Hough transform are often used to characterize the structural properties of objects within an image. Zhang & Couloigner (2006) expanded the work by also considering the shape characteristics
of feature. Four shape features are extracted namely, mean, compactness, eccentricity, and direction. The mean value should be larger for a parking lot pixel than a road/ bridge pixel. The compactness serves as a measure of how circular or elongated the object is within which the centre pixel is located which is useful in detecting building shapes. The eccentricity defines the location of the centre pixel in relation to the centroid of its parent object. Pixels on the border of an object will have greater values than those located near the centre which is one of the main supportive features in detecting gabled roof buildings. Road extraction by Jin & Davis (2005), building extraction by Lari, Z. et al. (2007) and bridge extraction by Loménie, N. (2003) are some of the works using geometrical classifiers.

2.3.1.7.4 Contextual classification

Contextual information is used by a number of feature extraction systems to guide the extraction process. Hinz et al. (2003) identify two types of contextual information for road detection as being global or local. Global contextual information is used to define regions having similar characteristics, such as curvature, width and colour. Local contextual information refers to the objects that are often found on or in close proximity to the structural features, such as cars, road markings, buildings or tree lanes. Identifying these structures could help to reinforce the existence of a wanted feature, particularly in instances where the region of interest is occluded. In the work done by Jin & Davis (2005) contextual information is combined with Structural and spectral information for improving the detection efficiency of building extraction. Bruzzone, L. & Carlin, L. (2006) proposed a method for global contextual classification based on edge densities within a given class. Four contextual classes are defined as urban, rural, montane, and hybrids of suburban and rural regions. Hinz et al. (2003) focus solely on urban areas and find that shadows, buildings, general occlusions, vehicles and convoys of vehicles provide the best local contextual information. Chaudhuri & Samal (2008) used contextual information for the extraction of bridges from multispectral images. The method first classifies the image into eight land-cover types using a majority-must-be-granted logic based on the multiseed supervised classification technique. The classified image is then categorized into a trilevel image: water, concrete, and background. Bridges are then recognized in this trilevel image by using a
knowledge-based approach that exploits the spatial arrangement of bridges and their surroundings using a five-step approach.

2.4 Summary

As it is evident from the topics discussed in this chapter, the process whereby structural features are extracted from remotely sensed imagery is elaborate and involves a number of techniques from numerous fields within computer vision and image processing. To develop a novel structural feature extraction system, one should have good idea in the field of remote sensing as the input images are remotely sensed data. This review serves only as an introduction to the basics of remote sensing considering satellite images alone and commonly used structural feature extraction techniques highlighting some of the more noteworthy works in the field. The following chapters will discuss each and every structural feature in more detail and analyze different methods developed.