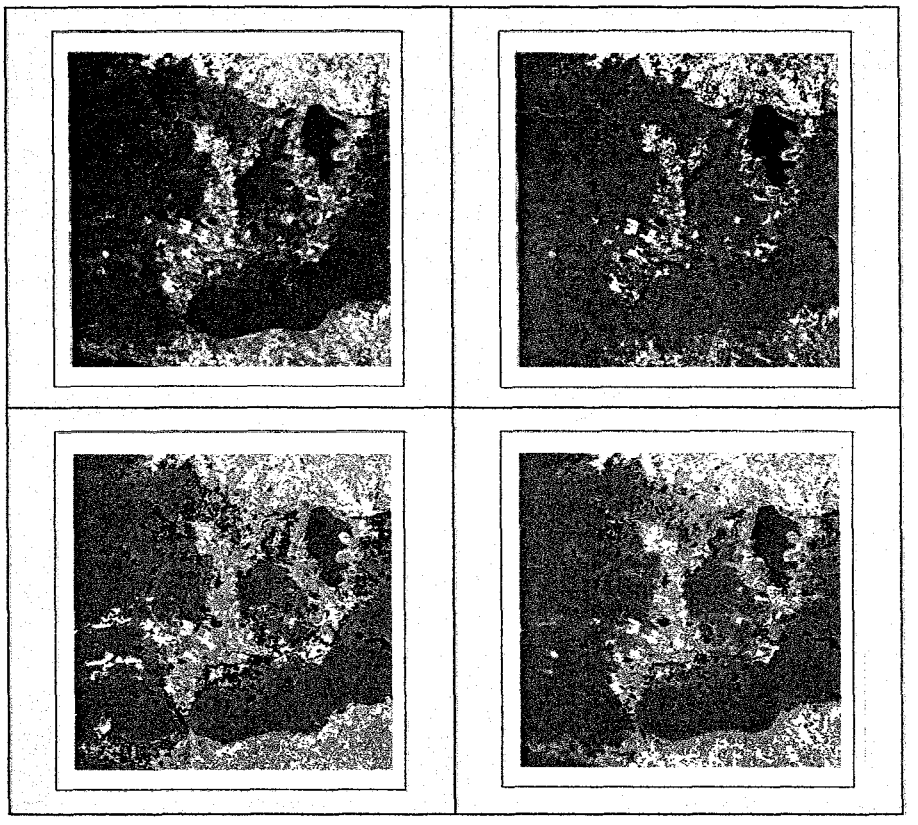
FUNDAMENTAL OF REMOTE SENSING AND DIGITAL IMAGE PROCESSING



DR. SUWIT ONGSOMWANG
SCHOOL OF REMOTE SENSING
INSTITUTE OF SCIENCE
SURANAREE UNIVERSITY OF TECHNOLOGY

2007

Preface

Digital image analysis and interpretation (106 602) is one of three main core subjects for Graduate students at School of Remote Sensing, Institute of Science, Suranaree University of Technology. This course provides general knowledge on digital image processing techniques to assist the interpretation and analysis of remotely-sensed data. Emphasis will be given on image enhancing technique using various standard image enhancing programs and the analysis of data from multispectral sensors. The main reference textbook used in the course based on John R. Jensen (2005) with entitled Introductory Digital Image Processing: A Remote Sensing Perspective. However, since the technology of remote sensing and techniques in digital image processing are rapidly change, the review and updating the knowledge in science of remote sensing and digital image processing are necessary. Therefore, the textbook with entitle "**Fundamental of Remote Sensing and Digital Image Processing**" are prepared based for the course (Digital image analysis and interpretation: 106 602). Major source of this textbooks were extracted from relevant remote sensing and digital image processing book included:

- Foody, G. M. 2004. Sub-Pixel Methods in Remote Sensing in In Remote Sensing Image Analysis: Including the Spatial Domain. De Jong, S.M., van der Meer, F.D. (eds.) Kluwer Academic Publishers, the Netherlands. pp: 37–49.
- Schowengerdt, R. A. 1997. Remote Sensing: Models and Methods for Image Processing. Academic Press, Inc. New York. 522 p.
- J Jensen, J. R. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd Edition. Practice Hall. 526 p.
- Lillesand, T. M. and R. W. Kiefer, and J. W. Chipman. 2004. Remote Sensing and Image Interpretation. John Wiley & Sons, Inc. New York, 763 pp.
- McCloy, K. R. (2006). Resources Management Information Systems: Remote Sensing, GIS and Modeling. 2Nd Edition. CRC Press Taylor & Francis Group, Fl, 575.

The main objectives of this textbook are as follows:

1. To provide concept and principle of remote sensing and sensors technology,
2. To review basic knowledge of aerial photographs and visual interpretation,
3. To provide fundamental and advance knowledge in digital image processing with digital change detection.

The content of the textbook that consists of nine chapters can be summarized in each chapter as following.

Chapter 1: Concepts and Fundamentals of Remote Sensing. This chapter is firstly reviewed definition of remote and the development of remote sensing. Then principle of remote sensing system is explained include (a) an ideal and real remote sensing systems, (b) electromagnetic remote sensing system, (c) electromagnetic radiation principle, (d) energy interactions in the atmosphere, (f) atmospheric windows, (h) interactions with the earth surface feature, and (i) radiometric concepts, terminology and units

Chapter 2: Sensor in Remote Sensing. This chapter provides the important information about aircraft and satellite remote sensing system include framing and scanning systems. In addition characteristics of selected remote sensor systems used for multispectral and hyperspectral data collection are summarized based on spectral, spatial, radiometric, and temporal resolutions.

Chapter 3: Aerial Photography and Visual Interpretation. Characteristics of aerial photographs consist of resolution, photographic scale and relief displacement are firstly summarized. The fundamentals of visual interpretation is then demonstrated in detail include (a) image interpretation task, (b) element of image interpretation, (c) image interpretation strategies, and (d) method of search.

Chapter 4: Satellite Data and Digital Image Processing. In this chapter, framework of digital image processing is introduced includes (a) characteristics of satellite data, (b) digital image resolution, (c) digital image processing, and (d) an ideal step of digital image processing. In addition, digital image processing systems are also summarized with summary of commercial and public software.

Chapter 5: Preprocessing in Digital Image Processing. Three main preprocessing operations include (1) image quality assessment and statistical evaluation, (2) radiometric correction and (3) geometric correction are demonstrated in detail in this chapter. Basic operations of image quality assessment and statistical evaluation are summarized include (a) histogram characteristics of remote sensor data (b) image metadata (c) viewing individual pixel brightness value (d) univariate descriptive image statistics,(e) multivariate descriptive image statistics and (f) feature space plots. While, principle of radiometric and geometric error in remote sensing system are explained in detail and basic and advance techniques for radiometric and geometric correction are discussed and demonstrated.

Chapter 6: Image Enhancement in Digital Image Processing. Selected image enhancement operations that have proven of value for visual analysis and/or subsequent digital image processing are reviewed and discussed in this chapter. They are three categories including (1) radiometric enhancement, (2) spatial enhancement and (3) spectral enhancement. Radiometric enhancement deals with the individual values of the pixels in the image consist of (a) linear

contrast enhancement and (b) nonlinear contrast stretch. While spatial enhancement modified pixel values based on the values of surrounding pixels include (a) spatial convolution filtering (b) Fourier transformation (c) CRISP (d) resolution merge (e) adaptive filter and (f) texture transform. Also, spectral enhancement modified more than one band of data spectral consist of (a) band ratioing (b) Principal Component Analysis (c) indices and (d) RGB to IHS transformation and back again.

Chapter 7: Image Classification in Digital Image Processing. Fundamental of image classification concept, logic and algorithm are here reviewed and discussed in this chapter. Selected common image classification algorithms that are applied for multispectral classification will be here separately explained in three groups: (1) parametric, (2) non-parametric and (3) nonmetric classification. Parametric classification includes (a) maximum likelihood classifiers and (b) clustering. While, nonparametric classification consists of (a) level-slice classifier (b) parallelepiped classifier (c) minimum distance to means classifier (d) nearest-neighbor classifiers (e) sub-pixel classification (f) artificial neural network (ANN) classifier. Also nonmetric classification consists of expert system.

Chapter 8: Accuracy Assessment. Review of source of error in remote sensing-derived thematic products is firstly reviewed and discussed. Then the fundamental of error matrix, sample size, sampling design and evaluation of error matrices explained. Three type of evaluation of error matrices include (a) descriptive evaluation of error matrices (b) discrete multivariate analytical technique: Kappa Analysis and (c) fuzzification of the error matrix.

Chapter 9: Digital Change Detection. In this chapter, principle of digital change detection is firstly described and then change detection algorithm are reviewed and explained. The digital change detection algorithm are here include (a) Write Function Memory Insertion (b) Multi-date Composite Image (c) Image Algebra (d) Post-classification Comparison (e) Binary Change Mask Applied to Date 2 (f) Ancillary Data Source as Date 1 (g) Spectral Change Vector Analysis (h) Chi-square Transformation (i) Cross-correlation and (j) Knowledge-based Vision Systems.

Dr. Suwit Ongsomwang
School of Remote Sensing
Institute of Science
Suranaree University of Technology
December 2007

Table of Contents

Preface ii

Chapter 1 Concepts and Fundamentals of Remote Sensing 1

1.1 Definition of Remote Sensing 1

1.2 The Development of Remote Sensing 1

1.3 An ideal and real remote sensing systems 8

1.4 Electromagnetic Remote Sensing System 10

1.5 Electromagnetic Radiation Principle 15

1.5.1 The Wave Model 15

1.5.2 The Particular Model 16

1.5.3 Stefan-Boltzmann' Law 17

1.5.4 Kirchhof's Law 19

1.5.5 Wien's Displacement Law 19

1.5.6 Planck's Law 19

1.6 Energy Interactions in the Atmosphere 21

1.6.1 Scattering 21

1.6.2 Absorption 23

1.6.3 Refraction 24

1.7 Atmospheric Windows 25

1.8 Interactions with the Earth Surface Feature 26

1.8.1 The interactions of electromagnetic radiation with vegetation 36

1.8.2 The interactions of electromagnetic radiation with water 38

1.8.3 The interactions of electromagnetic radiation with soil 40

1.9 Radiometric Concepts, Terminology and Units. 43

1.9.1 Radiometric Characteristics of Radiation Measurement 43

1.9.2 Radiometric Terminology and Units 45

Chapter 2 Sensor in Remote Sensing 49

2.1 Framing system 49

2. 2 Scanning system 50

2.2.1 Cross-Track Scanning System 50

2.2.2 Circular Scanning System 51

2.2.3 Along-Track Scanning System 51

2.2.4 Side Scanning System 52

Chapter 3 Aerial Photography and Visual Interpretation 71
3.1 Characteristics of Aerial Photographs 71
3.3.1 Resolution 71
3.3.2 Photographic Scale. 76
3.3.3 Relief Displacement. 76
3.2 Visual Interpretation 78
3.2.1 Image Interpretation Task 78
3.2.2 Element of image interpretation 80
3.2.3 Image Interpretation Strategies 86
3.2.4 Method of Search 87
Chapter 4 Satellite Data and Digital Image Processing 90
4.1 Characteristics of Satellite Data 90
4.2 Digital Image Resolution 95
4.3 Digital Image Processing 99
4.4 An Ideal Step of Digital Image Processing 108
Chapter 5 Preprocessing in Digital Image Processing 113
5.1 Image Quality Assessment and Statistical Evaluation 114
5.1.1 Histogram Characteristics of Remote Sensor Data 114
5.1.2 Image Metadata 115
5.1.3 Viewing Individual Pixel Brightness Value 117
5.1.4 Univariate Descriptive Image Statistics 118
5.1.5 Multivariate Descriptive Image Statistics 121
5.1.6 Feature Space Plots 123
5.2 Radiometric Correction 124
5.2.1 Source of image radiometry error 124
5.2.2 Type of Radiometric Correction 125
5.2.2.1 Sensor Calibration 125
5.2.2.2 Atmospheric Correction 129
5.2.2.3 Solar and Topographic Correction 143
5.3 Geometric Correction 146
5.3.1 Source of Image Geometry Error 146
5.3.1.1 Internal Geometric Errors 147
5.3.1.2 External Geometric Errors 153
5.3.2 Types of Geometric Correction 155

5.3.2.1 Image to Map Rectification 155
5.3.2.2 Image to Image Registration 156
5.3.2.3 Hybrid Approach to Image Rectification/Registration 156
5.3.3 Image to Map Geometric Rectification Logic 157
5.3.3.1 Spatial Interpolation 157
5.3.3.2 Intensity Interpolation 162
Chapter 6 Image Enhancement in Digital Image Processing 166
6.1 Radiometric Enhancement 167
6.1.1. Linear Contrast Enhancement 169
6.1.2 Nonlinear Contrast Stretch. 172
6.2 Spatial Enhancement 178
6.2.1 Spatial Convolution Filtering 179
6.2.1.1 Low-frequency Filtering in the Spatial Domain 181
6.2.1.2 High-frequency Filtering in the Spatial Domain 182
6.2.1.3 Edge Enhancement in the Spatial Domain 183
6.2.2 Fourier Transformation 190
6.2.2.1 Spatial Filter in Frequency Domain 194
6.2.3 CRISP 198
6.2.4 Resolution Merge 199
6.2.4.1 Principal Components Merge 199
6.2.4.2 Multiplicative Algorithm 200
6.2.4.3 Brovey Transform 200
6.2.5 Adaptive Filter 202
6.2.6 Texture Transform 203
6.3 Spectral Enhancement 210
6.3.1 Band Ratioing 211
6.3.2 Principal Component Analysis (PCA) 213
6.3.3 Indices 217
6.3.4 RGB to IHS Transformation and Back Again 221
Chapter 7 Image Classification in Digital Image Processing 224
7.1 Introduction 224
7.2 Classification Process 225
7.3 Land Use and Land Cover Classification Scheme 230
7.4 Training Site Selection and Feature selection 234

7.4.1 Training Site Selection 234
7.4.2 Feature selection 235
7.4.2.1 Graphic Methods of Feature Selection 235
7.4.2.2 Statistical Measures of Feature Selection 238
7.5 Image Classification Algorithm 240
7.5.1 Parametric Classification 240
7.5.1.1 Maximum Likelihood Classifiers 241
7.5.1.2 Clustering 246
7.5.2 Nonparametric Classification 256
7.5.2.1 Level-Slice Classifier 256
7.5.2.2 Parallelepiped Classifier 257
7.5.2.3 Minimum Distance to Means Classifier 260
7.5.2.4 Nearest-Neighbor Classifiers 264
7.5.2.5 Sub-Pixel Classification 265
7.5.2.6 Artificial Neural Network (ANN) Classifier 274
7.5.3 Nonmetric Classification 283
7.5.3.1 Expert Systems 283
Chapter 8 Accuracy Assessment 292
8.1 Source of Error in Remote Sensing-derived Thematic Products 292
8.2 General Steps to Assess the Accuracy of Thematic Information Derived from Remotely Sensed Data 294
8.2.1 The Error Matrix 296
8.2.2 Sample Size 297
8.2.2.1 Sample Size Based on Binomial Probability Theory 297
8.2.2.2 Sample Size Based on Multinomial Distribution 297
8.2.3 Sampling Design 299
8.2.3.1 Simple Random Sampling 299
8.2.3.2 Systematic Sampling 299
8.2.3.3 Stratified Random Sampling 300
8.2.3.4 Stratified Systematic Unaligned Sampling 300
8.2.3.5 Clustering Sampling 300
8.2.4 Evaluation of Error Matrices 302
8.2.4.1 Descriptive Evaluation of Error Matrices 302
8.2.4.2 Discrete Multivariate Analytical Technique: Kappa Analysis 303

8.2.4.3 Fuzzification of the Error Matrix 305
Chapter 9 Digital Change Detection 308
9.1 Step Required to Perform Change Detection 308
9.2 Remote sensor system considerations 311
9.3 Environmental Characteristics Considerations 312
9.4 Selection of a Change Detection Algorithm 316
9.4.1 Change Detection Using Write Function Memory Insertion 316
9.4.2 Multidate Composite Image Change Detection 318
9.4.3 Image Algebra Change Detection 319
9.4.4 Post-classification Comparison Change Detection 324
9.4.5 Change Detection Using a Binary Change Mask Applied to Date 2 326
9.4.6 Ancillary Data Source as Date 1 327
9.4.7 Spectral Change Vector Analysis 328
9.4.8 Chi-square Transformation Change Detection 330
9.4.9 Cross-correlation Change Detection 331
9.4.10 Knowledge-based Vision Systems for Detecting Change 332
9.4.11 Visual On-screen Detection and Digitization 333

List of Tables

Chapter 1 Concepts and Fundamentals of Remote Sensing 1
Table 1-1 Comparison of the two major time periods in remote sensing development 2
Table 1-2 Milestones in the history of remote sensing 7
Table 1-3 List of spectral regions in descending usefulness for monitoring green vegetation ... 38
Table 1-4 Radiometric Term, Symbol, Measure, Units 47
Chapter 2 Sensor in remote sensing 49
Table 2-1 Landsat MSS Landsat TM and Landsat ETM+ sensor system characteristics 57
Table 2-2 NOAA AVHRR sensor system characteristics 58
Table 2-3 Characteristics of the Sea-viewing Wide Field-of-View Sensor (SeaWiFS) 59
Table 2-4 System characteristics of the AMS and ATLAS 60
Table 2-5 SPOT HRV, SPOT HRVIR and SPOT Vegetation sensor system characteristics 61
Table 2-6 Indian Remote Sensing (IRS) satellite characteristics 62
Table 2-7 NASA ASTER sensor system characteristics 63
Table 2-8 Sensor characteristics of IKONOS, OrbView-3, and, QuickBird satellites 64
Table 2-9 THEOS sensor system characteristics. 65
Table 2-10 Leica Geostsystems Airborne Digital Sensors 40 (ADS-40) characteristics. 65

Table 2-11 Characteristics of AVIRIS and CASI hyperspectral remote sensing systems..... 66
Table 2-12 Characteristics of TERRA satellite: MODIS 67
Table 2-13 Leica Geosystems Emerge Digital Sensor System (DSS) characteristics 69
Table 2-14 Selected remote sensing systems and their characteristics. 70
Chapter 3 Aerial Photography and Visual Interpretation 71
Table 3-1 Minimum ground separation on typical aerial photographs 75
Table 3-2 Element of image interpretation 85
Table 3-3 Multidisciplinary scientist brings their unique training to the image interpretation process 89
Chapter 4 Satellite Data and Digital Image Processing 90
Table 4-1 Relationship between digitizer scanning spot size (IFOV) and the pixel ground resolution at various scales of aerial photography or image 94
Table 4-2 Image processing functions in quality digital image processing systems 102
Table 4-3 Major Functions of Selected Commercial Digital Image Processing Systems 104
Table 4-4 Major Functions of Selected Public Digital Image Processing Systems 107
Chapter 5 Preprocessing in Digital Image Processing 113
Table 5-1 Hypothetical Dataset of Brightness Values 120
Table 5-2 Variance-covariance matrix of the sample data 122
Table 5-3 Correlation Matrix of the Sample Data 123
Table 5-4 Pre-launch measurement of the TM calibration gain and offset coefficient 127
Table 5-5 Value of LMin and LMax for the Landsat MSS sensor system 129
Table 5-6 Example of atmospheric calibration techniques. 131
Table 5-7 Radiometric variables used in remote sensing 134
Table 5-8 Bilinear interpolation 163
Table 5-9 Cubic convolution interpolation 165
Chapter 6 Image Enhancement in Digital Image Processing 166
Table 6-1 Various kernels for Low-Frequency Filtering 182
Table 6-2 Various kernels for High-Frequency Filtering 183
Table 6-3 Various kernels for Linear Edge Enhancement 184
Table 6-4 Various kernels for Line Detection 187
Table 6-5 Catalog of local convolution filtering 187
Table 6-6 Selected remote sensing vegetation indices. 218

Chapter 7 **Image Classification in Digital Image Processing** 224
Table 7-1 USGS Land-Use/Land-Cover Classification System for Use with Remote Sensor Data 232
Table 7-2 Various distance measure methods for separability analysis 239
Table 7-3 Advantage and disadvantage of maximum likelihood decision rule 244
Table 7-4 Advantage and disadvantage of ISODATA decision rule 252
Table 7-5 Advantage and disadvantage of RGB clustering 254
Table 7-6 Advantage and disadvantage of parallelepiped decision rule 260
Table 7-7 Advantage and disadvantage of minimum distance decision rule 262
Table 7-8 The input values, activation and output value for simple two-channel ANN 278
Table 7-9 Advantage and disadvantage of ANN algorithm 282
Table 7-10 Examples of the use of expert system in remotely sensed data classification 283
Table 7-11 Hierarchical levels, classes and criteria 287
Table 7-12 A hypothesis (class), variables and conditions necessary to extract white fir 289
Table 7-13 Advantages and disadvantage of expert system 290
Chapter 8 **Accuracy Assessment** 292
Table 8-1 Characteristic of error matrix 296
Table8-2 Error matrix and their accuracy assessment by simple descriptive statistics 302
Table 8-3 Error matrix and their accuracy assessment by Kappa Analysis 304
Table 8-4 Error matrix and their accuracy assessment with Fuzzy Logic Rule 306
Table 8-5 Error matrix and their accuracy assessment without Fuzzy Logic Rule 307
Chapter 9 **Digital Change Detection** 308
Table 9-1 Sector code definitions for change vector analysis using three bands 329

List of Figures

Chapter 1 **Concepts and Fundamentals of Remote Sensing** 1
Figure 1-1 Components of an ideal remote sensing system 9
Figure 1-2 Electromagnetic remote sensing system of earth resource 11
Figure 1-3 A conceptual view of an Earth observational system 12
Figure 1-4 Remote sensing process 14
Figure 1-5 Types of energy transfer: conduction, convection and radiation 15
Figure 1-6 An electromagnetic wave 16
Figure 1-7 Electromagnetic spectrum diagrams 18
Figure 1-8 Spectral distribution of energy radiated from blackbody of various temperatures 20
Figure 1-9 Interrelationship between temperature, wavelength, frequency radiant energy,

radiant exitance and the point of maximum spectral radiant exitance..... 20
Figure 1-10 Atmospheric scattering 22
Figure 1-11 Atmospheric absorption by various gases 24
Figure 1-12 Atmospheric refraction..... 25
Figure 1-13 Spectral characteristics of energy sources, atmospheric effects, and sensing system 26
Figure 1-14 Basic interactions between electromagnetic energy and the earth surface feature... 27
Figure 1-15 Typical spectral reflectance curve for vegetation, soil and water..... 28
Figure 1-16 The nature of specular and diffuse reflectance..... 29
Figure 1-17 Hemispherical, directional and bi-directional reflectance 30
Figure 1-18 The bi-directional reflectance function (BRDF) of a recent ploughed bare field at 25° solar elevation for 662 nm wavelength 31
Figure 1-19 The bi-directional reflectance function (BRDF) of a sand shinnery oak rangelands community at 31° solar zenith angle at = 662 nm 31
Figure 1-20 The bi-directional reflectance function (BRDF) of a sand shinnery oak rangelands community at 31° solar zenith angle at = 862 nm 32
Figure 1-21 The bi-directional reflectance effects..... 32
Figure 1-22 Atmospheric effects influencing the measurement of reflected solar energy 34
Figure 1-23 Effects of seasonal change on solar elevation angle 35
Figure 1-24 Significant spectral responses characteristics of green vegetation..... 37
Figure 1-25 Water bodies receive irradiance from the Sun and atmosphere 39
Figure 1-26 The absorption characteristics of pure water..... 40
Figure 1-27 The reflectance characteristic of variant water..... 40
Figure 1-28 Soils and rocks receive irradiance from the Sun and atmosphere..... 41
Figure 1-29 Representative spectra for the reflectance of soils..... 43
Figure 1-30 Reflectance of a typical soil with changes in moisture content..... 43
Figure 1-31 Measurement of incoming and outgoing radiation 44
Figure 1-32 Projected area cosine in a viewing direction other than normal 44
Figure 1-33 Defining angle used in radiation measurement..... 45
Figure 1-34 Relationship among the various terms in hemispherical and directional radiation measurement..... 46
Figure 1-35 Characteristics of radiant flux density..... 48
Figure 1-36 Concept of radiance 48

Chapter 2 Sensor in remote sensing 49
Figure 2-1 Framing system to acquiring remote sensing images 49
Figure 2-2 Cross-Track Scanner 50
Figure 2-3 Circular Scanner 51
Figure 2-4 Along-Track Scanner 52
Figure 2-5 Side Scanning System 53
Figure 2-6 Six types of remote sensor systems 54
Chapter 3 Aerial Photography and Visual Interpretation 71
Figure 3-1 Resolution and detection targets with high contrast ratio 73
Figure 3-2 Resolution and detection targets with low contrast ratio 74
Figure 3-3 Ground resolution and minimum ground resolution on aerial photographs 75
Figure 3-4 Geometry of relief displacement on a vertical aerial photograph 77
Figure 3-5 Element of Image Interpretation – Tone 81
Figure 3-6 Element of Image Interpretation – Texture 82
Figure 3-7 Element of Image Interpretation – Shadow 82
Figure 3-8 Element of Image Interpretation – Texture 83
Figure 3-9 Element of Image Interpretation – Site, Situation and Association 83
Figure 3-10 Element of Image Interpretation – Shape 84
Figure 3-11 Element of Image Interpretation – Size 84
Chapter 4 Satellite Data and Digital Image Processing 90
Figure 4-1 Digital remote sensor data 90
Figure 4-2 Multispectral concept and data representation 91
Figure 4-3 Analog-to-digital conversion process 91
Figure 4-4 Schematic of flatbed densitometer 92
Figure 4-5 Schematic of drum densitometer 92
Figure 4-6 Video densitometer 93
Figure 4-7 Video densitometer 93
Figure 4-8 Characteristics of spectral resolution 95
Figure 4-9 Characteristics of spatial resolution 96
Figure 4-10 Characteristics of radiometric resolution 97
Figure 4-11 Characteristics of temporal resolution based on Landsat-5 97
Figure 4-12 Four types of resolution of Landsat TM-Band 2 98
Figure 4-13 Spatial and temporal resolution considerations for selected application 98
Figure 4-14 A typical digital image processing laboratory 100

Figure 4-15 Idealized sequence for digital image analysis 108
Figure 4-16 Analog (visual) and digital image processing image processing of remotely sensed data use the fundamental elements of image processing 112
Chapter 5 Preprocessing in Digital Image Processing 113
Figure 5-1 Histograms of symmetric and skewed distribution 115
Figure 5-2 Histogram of a single band of Landsat Thematic Mapper data. 116
Figure 5-3 Cursor evaluation of individual pixel brightness values 117
Figure 5-4 The areas under the normal curve for various standard deviations 119
Figure 5-5 Feature Space Plot of Landsat-TM band 3 and 4 124
Figure 5-6 Data flow for calibration of remote sensing images to physical units 126
Figure 5-7 Radiometric response function for an individual TM channel 128
Figure 5-8 Inverse of radiometric response function for an individual TM channel 128
Figure 5-9 Various paths of radiance received by a remote sensing system 133
Figure 5-10 An example absolute atmospheric correction using ATCOR 138
Figure 5-11 Absolute Atmospheric Correction Using Empirical Line Calibration 139
Figure 5-12 Single-image Normalization Using Histogram Adjustment 141
Figure 5-13 Multiple-date Image Normalization Using Regression 143
Figure 5-14 Representation of the Sun's angle of the incidence and solar zenith angle 144
Figure 5-15 Image offset (Deskew) caused by earth rotation effect 149
Figure 5-16 The geometrically configuration of scanning system-induced variation in ground resolution cell size 150
Figure 5-17 One-dimensional relief displacement and tangential scale distortion in scanning system 152
Figure 5-18 Geometric modification of remotely sensed data caused by change in platform altitude and attitude 154
Figure 5-19 Example of image-to-map rectification 156
Figure 5-20 Example of image-to-image hybrid registration 157
Figure 5-21 Concept of how different-order transformations fit a hypothetical surface 158
Figure 5-22 The logic of filling a rectified output matrix with values from an unrectified input image matrix 160
Figure 5-23 Linear transformation 161
Figure 5-24 Non-linear transformation 161
Figure 5-25 Spatial interpolation algorithm 164

Chapter 6 Image Enhancement in Digital Image Processing 166
Figure 6-1 Histograms of radiometrically enhanced dat 168
Figure 6-2 Graph of a Lookup Table 168
Figure 6-3: Enhancement with Lookup Tables 169
Figure 6-4 Minimum-Maximum Contrast Stretching 170
Figure 6-5 Percentage Linear and Standard Deviation Contrast Stretching 171
Figure 6-6 Piecewise Linear Contrast Stretch 171
Figure 6-7 Nonlinear radiometric enhancement 172
Figure 6-8 Histogram Equalization 173
Figure 6-9 Original Histogram Data 175
Figure 6-10 Equalized Histogram Data 175
Figure 6-11 Comparison of radiometric enhancement 176
Figure 6-12 Example of Brightness Inversion 177
Figure 6-13 Spatial Frequencies 178
Figure 6-14 Applying a Convolution Kernel 179
Figure 6-15 Application of various linear enhancement to Landsat TM 188
Figure 6-16 Application of various non linear enhancement to Landsat TM 189
Figure 6-17 One-Dimensional Fourier Analysis 190
Figure 6-18 Application of Fourier transform to the three different sub-images 192
Figure 6-19 Two examples of stationary periodic noise and their Fourier transforms 193
Figure 6-20 A filtering algorithm that uses the Fourier transform to compute a spatial domain convolution 195
Figure 6-21 Spatial Filter in Frequency Domain using a Fourier transform 196
Figure 6-22 Application of Fourier transform to a portion of Landsat TM 197
Figure 6-23 Application of Crisp Filtering to Landsat TM 198
Figure 6-24 Resolution merge between SPOT and Landsat TM 201
Figure 6-25 Application of Adaptive Filtering to Landsat TM 202
Figure 6-26 The eight nearest neighbors of pixel X according to angle ϕ used in the creator of spatial-dependency matrices for the measurement of image texture 205
Figure 6-27 A 5 X 5 window of pixels and their image values 208
Figure 6-28 The ratio of various Landsat TM bands 212
Figure 6-29 Principal Component Analysis 213
Figure 6-30 Distribution of total image variance across the original spectral bands and across the principal components 215

Figure 6-31 Scene dependence in the percentage of total TM image variance captured by each eigenvalue. The TIR band is excluded 216
Figure 6-32 Six principal component images of Landsat-TM 216
Figure 6-33 Various vegetation indices of Landsat TM 220
Figure 6-34 Intensity-hue-saturation (HIS) color coordinate system 222
Figure 6-35 Relationship between intensity-hue-saturation (HIS) color coordinate system and RGB coordinate system 222
Figure 6-36 Resolution merge between SPOT and Landsat TM data based on modified IHS resolution merge 223
Chapter 7 Image Classification in Digital Image Processing 224
Figure 7-1 The data flow in a classification process 226
Figure 7-2 Comparison between a traditional hard and soft classification logic 230
Figure 7-3 Relationship between the level of detail required and the spatial resolution of representative remote sensing system for vegetation inventories 233
Figure 7-4 Coincident spectral plots for training data obtained in five bands for six cover type 236
Figure 7-5 Cospectral mean vector plot of 49 clusters 237
Figure 7-6 Ellipse Evaluation of Signatures 237
Figure 7-7 Simple parallelepiped displayed in pseudo three-dimensional space 238
Figure 7-8 How the maximum likelihood decision rule function 244
Figure 7-9 Land use and land cover classification using maximum likelihood decision rule 245
Figure 7-10 An idealized data distribution during three iterations of the K-means clustering algorithms with the nearest-mean decision criterion 247
Figure 7-11 Typical behavior of the net mean migration from one iteration to the next in the K-mean algorithm 248
Figure 7-12 An idealized data distribution during initial stage up to n iterations of the ISODATA clustering algorithms 250
Figure 7-13 Land use and land cover classification using ISODATA clustering algorithms 251
Figure 7-14 RGB clustering algorithm 254
Figure 7-15 Land use and land cover classification using RGB clustering algorithms 255
Figure 7-16 Level-slice decision boundaries for three classes in two dimensions 257
Figure 7-17 Parallelepiped classification using ± two Standard deviations as limits 258
Figure 7-18 Land use and land cover classification using parallelepiped decision rule 259
Figure 7-19 Calculation of spectral distance to means 261

Figure 7-20 Distance use in a minimum distance to means classification algorithm 262
Figure 7-21 Land use and land cover classification by using Minimum distance decision rule ... 263
Figure 7-22 Hypothetical example of nearest-neighbor classification 265
Figure 7-23 Some common origins of mixed pixel problems 266
Figure 7-24 The linear mixing model for a single GIFOV 267
Figure 7-25 Three possible choices for endmembers for three classes 267
Figure 7-26 An example of a typical sub-pixel analysis 269
Figure 7-27 The logic of fuzzy classification 270
Figure 7-28 Basic structure of a three layer Artificial Neural Network 275
Figure 7-29 The components of a processing element 276
Figure 7-30 The sigmoid activation function 276
Figure 7-31 Mathematic model of neuron 277
Figure 7-32 Two thresholding functions used to transform the activation to output values at a synapse 278
Figure 7-33 The data space and the decision surface for a simple two channel, single hidden . 279
Figure 7-34 The Components of a Typical Rule-based Expert System 284
Figure 7-35 Example of a Decision Tree Branch 287
Figure 7-36 Split Rule Decision Tree Branch 288
Figure 7-37 A human-derived decision-tree expert system with a rule and conditions 289
Figure 7-38 Land use and land cover classification using expert classification 291
Chapter 8 Accuracy Assessment 292
Figure 8-1 Source of Error in Remote Sensing-derived Thematic Products 294
Figure 8-2 General steps to assess the accuracy of thematic information derived from remotely sensed data 295
Figure 8-3 Geographic Sampling methods 301
Chapter 9 Digital Change Detection 308
Figure 9-1 General data processing element for a remote sensing change detection application project or scientific study 309
Figure 9-2 Sequential steps outlining the project formulation process for remote sensing change detection project 309
Figure 9-3 The general steps used to perform digital change detection of remotely sensed data 310
Figure 9-4 Phenological cycle of Cattails and Waterlilies in Par Pond in SC 314

Figure 9-5 Algorithm, advantage and disadvantage of write function memory insertion change detection 317
Figure 9-6 An example of change detection using write function memory insertion 317
Figure 9-7 Algorithm, advantage and disadvantage of multiple-date composite image change detection 318
Figure 9-8 Principal components derived from a multiple-date dataset of Landsat TM in 1999 and 2004 319
Figure 9-9 Algorithm, advantage and disadvantage of image algebra change detection 320
Figure 9-10 Image differencing change detection Scaling alternatives and placement of user-specified thresholds in the change image histogram 321
Figure 9-11 Image differencing change detection 322
Figure 9-12 Vegetation cover change detection based on differencing NDVI. 323
Figure 9-13 Algorithm, advantage and disadvantage of post-classification comparison change detection 324
Figure 9-14 Land cover change 2001-2004 325
Figure 9-15 Algorithm, advantage and disadvantage of change detection using a binary change mask applied to date 2 326
Figure 9-16 Algorithm, advantage and disadvantage of change detection using an ancillary as date 1 327
Figure 9-17 Schematic diagram of the spectral change detection method 329
Figure 9-18 Possible change sector codes for a pixel measured in three bands on two dates ... 330
Figure 9-19 Algorithm, advantage and disadvantage of cross-correlation change detection 332
Figure 9-20 Visual on-screen change detection in Tsunami's effect 333

Chapter 1: Concepts and Fundamentals of Remote Sensing

1.1 Definition of Remote Sensing

Linz and Simonett (1976) defined remote sensing as the acquisition data of an object without touch or contact.

Lillesand and Kiefer (1979) defined remote sensing as 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 contact with the object, area, or phenomenon under investigation.

Barrett and Curtis (1982) defined remote sensing as the observation of a target by a device some distance away from it.

Colwell (1983) defined remote sensing as the measurement or acquisition of information of some property of an object or phenomena, by a recording device that is not in physical or intimate contact with the objective or phenomenon under study.

Curran (1985) defined remote sensing as the use of electromagnetic radiation sensors to record images of the environment which can be interpreted to yield useful information.

Sabins (1987) broadly defined remote sensing as collecting and interpreting information about target without being in physical contact with object.

Campbell (1987) defined remote sensing as the science of deriving information about the earth's land water areas from images acquired at a distance.

Colwell (1997) defined remote sensing as the art, science, and technology of obtaining reliable information about physical objects and the environment, through the process of recording, measuring and interpreting imagery and digital representations of energy patterns derived from non-contact sensor system.

In conclusion, remote sensing can be defined as the art, science, and technology of obtaining reliable information about object, area, and phenomenon of earth's resources and environment, through the process of recording, measuring and interpreting imagery and digital representations of energy patterns derived from non-contact sensor system.

1.2 The Development of Remote Sensing

Simonett (1983) reported that Colwell (1979) divided the development of remote sensing into two general areas. Prior about 1960, aerial photograph was the sole system used in remote sensing. With the advent of space program in the early 1960's and the first photographs from the Mercury, the pace of technological development for remote sensing accelerated. Table 1-1 shows a comparison of the two time periods.

Table 1-1: Comparison of the two major time periods in remote sensing development (Colwell, 1979)

Prior to the space age (1860-1960)	Since 1960.
A. Only one kind and date of photography.	A. Many kinds and dates of remote sensing data.
B. Heavy reliance on the human analysis of unenhance images.	B. Heavy reliance on the machine analysis and enhancement.
C. Extensive use of photo interpretation keys.	C. Minimal use of photo interpretation keys.
D. Relatively good military/civil relations with respect to remote sensing.	D. Relatively poor military/civil relations with respect to remote sensing.
E. Few problems with uniformed opportunities.	E. Many problem uninformed opportunities.
F. Minimal applicability of the "multi" concept.	F. Extensive applicability of the "multi" concept.
G. Equipment simple and inexpensive; readily operated and maintained by resource-oriented workers.	G. Equipment complex and expensive; not readily operated and maintained by resource-oriented workers.
H. Little concern about the renewability of resources, environmental protection, global resource information system, and associated problems related to "signature extension", "complexity of an area's structure", and/or the threat imposed by "economic weaponry".	H. Much concern about the renewability of resource, environmental protection, global resource information system, and associated problems related to "signature extension", "complexity of an area's structure", and/or the threat imposed by "economic weaponry.
I. Heavy resistance to "technology acceptance" by potential users of remote sensing-device information.	I. Continuing heavy resistance to "technology acceptance" by potential users of remote sensing-derived information.

From Simonett (1983)

Elachi (1987) had also confined the development of remote sensing and summarized as following:

The early development of remote sensing as a scientific field was clearly tied to developments in photography. The first photographs were reportedly taken by Daguerre and Niepce in 1839. The following year, Arago, Director of the Paris Observatory, advocated the use of photography for topographic purposes. In 1849 Colonel Aime' Laussedat, an officer in the French Corps. of Engineerings, embarked on an exhaustive program to use photography in topographic mapping. In 1858 balloons were being used to acquire photography of large areas. This was followed by the use of kites in the 1880s and pigeons in the early 1900s to carry cameras to many hundreds meters of altitude. The advent of the airplane made aerial photography a very useful tool because acquisition of data specific areas and under controlled conditions became possible. The first recorded photographs were taken from an airplane piloted by Wilbur Wright in 1890 over Centocelli, Italy.

Color photography became available in the mid 1930s. At the same time work was continuing on the development of films that were sensitive to near-infrared radiation. Near-infrared photography was particularly useful for haze penetration. During World War II research was conducted on the spectral reflectance properties of natural terrain and the availability of photographic emulsions for aerial color infrared photography. The main incentive was to develop techniques for camouflage detection.

In 1956, Colwell performed some of the early experiments on the use of special purpose aerial photography for the classification and recognition of vegetation types and the detection of diseased and damage vegetation. Beginning in the mid-1969s, a large and multispectral photography were undertaken under the sponsorships of NASA, leading to the launch of multispectral imageries on the Landsat satellites in the 1970s.

As the long wavelength end of the spectrum, active microwave systems have been used since early this century and particularly after World War II to detect and track moving objects such as ships and later, planes. More recently, active microwave sensors have been developed providing two-dimensional images that look very similar to regular photography, except the image brightness is a reflection of the scattering properties of the surface in the microwave region. Passive microwave sensors were also developed to provide photographs of the microwave emission of natural objects.

The tracking and ranging capabilities of radio systems were known as early as 1889, when Heirich Hertz showed that solid objects reflected radio waves. In the first quarter of this century, a number of investigations were conducted in the of radar systems for the detection and tracking of

ships and planes and for the study of the ionosphere. Radar work expanded dramatically during World War II. Today, the diversity of applications for radar is truly starting. It is being used to study ocean surfaces, lower and upper atmospheric phenomena, subsurface and surface land structures, and surface cover. Radar sensors exist in many different configurations. These include altimeters to provide topographic measurement, scatterometers to measure surface roughness, and imageries.

In the mid-1950s extensive work took place in the development of real aperture airborne imaging radars. At about the same time, work was on going in developing synthetic aperture imaging radar (SAR), which use coherent signals to achieve to the scientific community in the mid-1960s. Since then, work has continued at a number of institutions to develop the capacity of radar sensors to study natural surface. This work led to the orbital flight around the Earth of the SEASAT SAR (1987) and The Shuttle Imaging Radar (1981, 1984).

The most recently introduced remote sensing instrument is the laser, which was first developed in 1960. It is mainly being used for atmospheric studies, topographic mapping, and surface studies by fluorescence.

In addition, de Jong et al. (2004) reviewed the historic development of remote sensing in each decade as follows:

In 1859 Gaspard Tournachon took an oblique photograph of a small village near Paris from a balloon. With this picture the era of earth observation and remote sensing had started. Other people all over the world soon followed his example. During the Civil War in the United States aerial photography from balloons played an important role to reveal the defence positions in Virginia. Likewise other scientific and technical developments this Civil War time in the United States speeded up the development of photography, lenses and applied airborne use of this technology. Although the space era of remote sensing was still far away after the Civil war, already in 1891 patents were granted in Germany to successful designs of rockets with imaging systems under the title: 'new or improved apparatus for obtaining bird's eye photographic views of the earth'. The design comprised a rocket propelled camera system that was recovered by a parachute.

The next period of fast developments in earth observation took place in Europe and not in the United States. It was during World War I that airplanes were used on a large scale for photoreconnaissance. Aircrafts proved to be more reliable and more stable platforms for earth observations than balloons. In the period between World War I and World War II a start was made with the civilian use of aerial photos. Application fields of airborne photos included at that time geology, forestry, agriculture and cartography. These developments lead to improved cameras, films and interpretation equipment. The most important developments of aerial photography and

photo interpretation took place during World War II. During this time span the development of other imaging systems such as near-infrared photography, thermal sensing and radar took place. Near-infrared photography and thermal infrared proved very valuable to separate real vegetation from camouflage. The first successful airborne imaging radar was not used for civilian purposes but proved valuable for nighttime bombing. As such the system was called by the military: 'plan position indicator' and was developed in Great Britain in 1941.

After the wars in the 1950s remote sensing systems continued to evolve from the systems developed for war efforts. Color infrared photography (CIR) was found to be of great use for the plant sciences. In 1956 Colwell conducted experiments on the use of CIR for the classification and recognition of vegetation types and the detection of diseased and damaged or stressed vegetation. It was also in the 1950s that significant progress in radar technology was achieved. Two types of radar were developed at that time: SLAR: side-looking airborne radar and SAR: Synthetic Aperture Radar. Either development aimed at the acquisition of images at the highest possible resolution. Crucial to the SAR development was the ability to finely resolve the Doppler frequencies using a frequency analyses algorithm on the returning radar signal by the US Air Force research centre.

In the early 1960s the US started placing remote sensors in space for weather observation and later for land observations. TIROS (Television Infrared Observation Satellite) was the first meteorological satellite. A long series of meteorological satellites followed this one. 1960 was also the beginning of a famous US military space imaging reconnaissance program called Corona. Unfortunately, much of this program remained classified until 1995. In 1970 the TIROS program was renamed into NOAA (National Oceanic and Atmospheric Administration). Until today the NOAA Advanced Very High Resolution Radiometer (AVHRR) is orbiting the globe and collecting information on weather patterns in visible, near infrared and thermal wavelengths. NOAA-17 was launched on June 24, 2002. The 1950s and 1960s were also important for the organizational development of remote sensing. Various civil research organizations and universities became highly interested in these new technologies. This resulted in the start of various professional organizations and the publishing of remote sensing journals such as the IEEE Transactions on Geoscience and Remote Sensing, International Journal of Remote Sensing, Remote Sensing of Environment and Photogrammetric Engineering & Remote Sensing. Today remote sensing is not only taught at the university level but also at high schools.

In the early 70s the first satellite specifically designed to collect data of the earth's surface and its resources was developed and launched: ERTS-1 Earth Resources Technology Satellite. Later, in 1975, this program was renamed into Landsat. This first earth resources satellite was in fact a modified Nimbus weather satellite carrying two types of sensors: a four waveband multi-

spectral scanner (MSS) and three return beam vidicon television cameras (RBV). The sensors aboard this satellite proved to be able to collect high quality images at a reasonable spatial resolution. These images gave remote sensing a worldwide recognition as a valuable technology. The main advantages recognized at that time were: ready availability of images for most of the world, lack of political, security and copyright restrictions, low cost, repetitive multi-spectral coverage and minimal image distortion. Landsat 2 and 3 were launched in 1975 and 1978, respectively, and carried the same payload as the first satellite of this series. The payload was changed in 1982 with Landsat 4. The technically more advanced Thematic Mapper (TM) sensor replaced the RBV. An improved design of the TM, the ETM+ (Enhanced Thematic Mapper) was mounted aboard Landsat 7 and launched in 1999. The Landsat series is a very successful program, various MSS and TM sensors exceeded by far its design life time and its imagery is probably the most widely used data in the Earth sciences. One black spot on its history record is the 'failure upon launch' of Landsat 6 in 1993.

Various other successful earth observation missions carried out by other countries followed the Landsat program. In 1978 the French government decided to develop their own earth observation program. This program resulted in the launch of the first SPOT satellite in 1986. To the original SPOT design of three spectral bands a new sensor called Vegetation was added aboard SPOT-4 in 1998. Other earth observation missions are the Indian Remote Sensing Program (IRS) started in 1988, the Russian Resurs series first launched in 1985 and the Japanese ADEOS (Advanced Earth Observing Satellite) put in orbit in 1996. The European Space Agency (ESA) launched its first remote sensing satellite, ERS-1, in the year 1991. ERS carries various types of sensors aboard among which the AMI, a C-band (5 cm radar) active microwave instrument. The main focus of the ERS program is oceanographic applications although it is also widely used for monitoring tropical forests. In 1995 ERS-2 was successfully launched. In March 2002 ESA launched Envisat-1, an earth observation satellite with an impressive payload of 13 instruments such as a synthetic aperture radar (ASAR) and a Medium Resolution Imaging Spectrometer (MERIS). An important recent development is the launch of high-resolution earth observation systems such as IKONOS and QuickBird. These systems have multi-spectral systems collecting information in 4 bands (blue, green, red and near-infrared) at a spatial resolution of 4 meters or better. IKONOS has also a panchromatic mode (0.45-0.90 μm) with a spatial resolution of 1 m. With IKONOS, QuickBird and similar systems, space borne remote sensing approaches the quality of airborne photography. Table 1-2 shows a few important dates in the development of remote sensing.

Table 1-2: Milestones in the history of remote sensing

1800	Discovery of Infrared by Sir W. Herschel
1839	Beginning of Practice of Photography
1847	Infrared Spectrum Shown by J.B.L. Foucault
1859	Photography from balloons
1873	Theory of Electromagnetic Spectrum by J.C. Maxwell
1909	Photography from Airplanes
1916	World War I: Aerial Reconnaissance
1935	Development of Radar in Germany
1940	WW II: Applications of Non-Visible Part of EMS
1950	Military Research and Development
1959	First Space Photograph of the Earth (Explorer-6)
1960	First TIROS Meteorological Satellite Launched
1970	Skylab Remote Sensing Observations from Space
1971	Launch of Landsat-1 (ERTS-1): MSS sensor
1972	Rapid Advances in digital image processing
1978	Launch of Seasat (first spaceborne L-band radar)
1982	Launch of Landsat-4: new Generation of Landsat sensors TM
1986	French Commercial Earth Observation Satellite SPOT
1986	Development Hyperspectral Sensors
1990	Development High Resolution Spaceborne Systems
1990	First Commercial Developments in Remote Sensing
1991	Launch of the first European Remote Sensing Satellite ERS1 (active radar)
1998	Towards Cheap One-Goal Satellite Missions
1999	Launch of EOS-TERRA: NASA Earth Observing Mission
1999	Launch of IKONOS, very high spatial resolution sensor system
2001	Launch of Landsat-7 with new ETM+ sensor
2001	Launch of QuickBird, very high spatial resolution sensor system
2002	Launch of ESA's Envisat with 10 advanced instruments

From: de Jong et al, 2004

1.3 An ideal and real remote sensing systems

Lillesand and Kiefer (1979) introduced the basic components of an ideal remote sensing system (Figure 1-1) and explained the real remote sensing system. The ideal remote sensing system included:

1. A Uniform Energy Source. This source would provide energy over all wavelength, at a constant, known, high level of output, irrespective of time and place.

2. A Noninterfering Atmosphere. This would be an atmosphere that would not modify the energy from the source in any manner, whether that energy were on its way to the earth's surface or coming from it. Again ideally, this would hold irrespective of wavelength, time, place, and sensing altitude involved.

3. A Series of Unique Energy/Matter Interactions at the Earth's Surface. These interactions would generate reflected and/or emitted signals that are not only selective with respect to wavelength, but also are known, invariant, and unique to each and every earth surface feature type and subtype of interest.

4. A Super Sensor. This would be a sensor highly sensitive to all wavelengths, yielding spatially detailed data on the absolute brightness (or radiance) from a scene as a function of wavelength, throughout the spectrum. This super sensor would be simple, reliable, require virtually no power or space, be accurate, and economical to operate.

5. Real-Time Data Handling System. In this system, the instant the radiance versus wavelength response over a terrain element were generated, it would be processed into an interpretable format and recognized as being unique to the particular terrain element from where it came. This process would be performed nearly instantaneous (real time), providing timely information.

6. Multiple Data Users. These people would have knowledge of great depth, both their respective disciplines and of remote sensing data acquisition and analysis techniques. The same set of data would become various forms of information for different users, because of their wealth of knowledge about the particular earth resource being sensed. With this information, the various users would make profound, wise decisions about how best to manage the earth resource under scrutiny and these management decisions would be implemented.

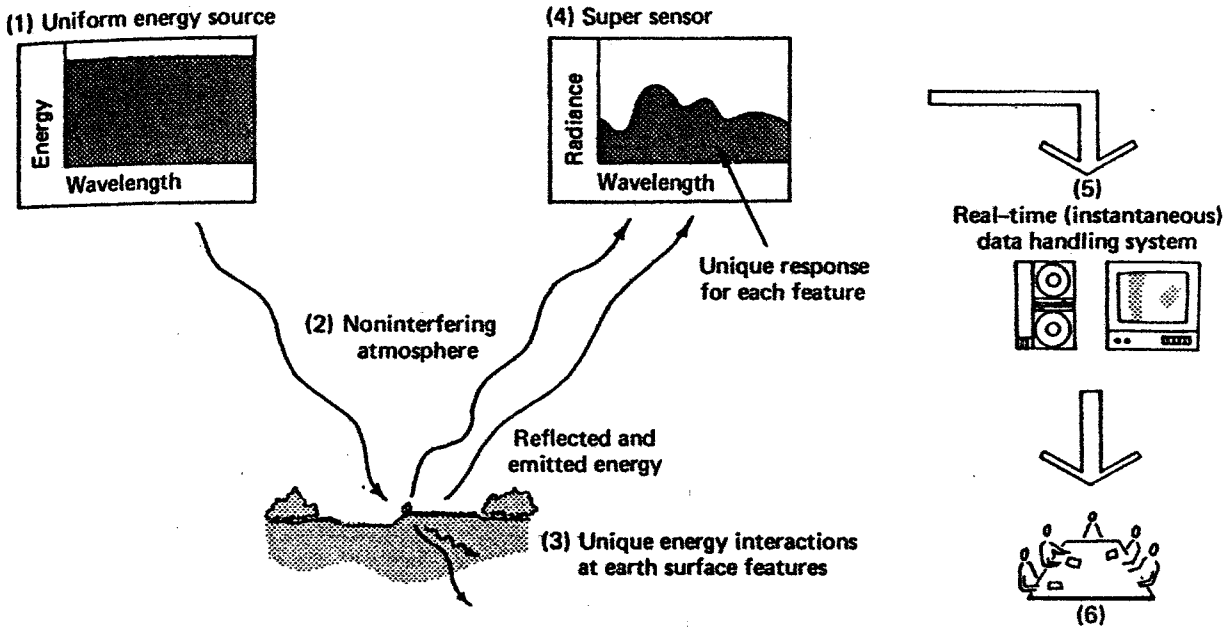


Figure 1-1: Components of an ideal remote sensing system (From Lillesand and Kiefer, 1979)

Unfortunately, an ideal remote sensing system as described above does not exist (Lillesand and Kiefer, 1979). Regarding the elements of the ideal system that has been suggested, the following general shortcomings of real systems should be recognized:

1. The Energy Source. All passive remote sensing systems rely on energy that is either reflected and/or emitted from earth surface features. Solar energy levels obviously vary with respect to time and location, and different earth surface materials emit energy to varying degrees of efficiency. While we have some control over the nature of sources of energy for active systems, the sources of energy used in all real systems are generally non-uniform with respect to wavelength and their properties vary with time and locations. Consequently, we normally must calibrate for source characteristics on a mission-by-mission basis, or deal with relative energy units sensed at any given time and location.

2. The Atmosphere. The atmosphere normally compounds the problems introduced by energy source variation. To some extent, the atmosphere always modifies the strength and spectral distribution of the energy received by a sensor. It restricts "where we can look" spectrally and its effects vary with wavelength, time and place. Elimination of, or compensation for, atmospheric effects via some form of calibration is particularly important in these applications where repetitive observations of the same geographical area are involved.

3. The Energy/Matter Interactions at the Earth's Surface. Remote sensing would be simple if each and every material reflected and/or emitted energy in a unique known way. Although spectral signatures play a central role in detecting, identifying, and analyzing earth surface materials, the spectral world is full of ambiguity. Radically different material types can have great spectral similarity, making differentiation difficult. Furthermore, the general understanding of the energy/matter interactions for earth surfaces is at an elementary level for some materials and virtually nonexistent for others.

4. The Sensor. No single sensor is sensitive to all wavelengths. All real sensors have fixed limits of spectral sensitivity. They also have a limit on how small an object on the earth's surface can be and still be seen by a sensor as being separate from its surrounding. This limit, called the spatial resolution of a sensor, is an indication of how well a sensor can record spatial detail.

5. The Data Handling System. The capability of current remote sensors to generate data for exceeds the current capacity to handle these data. This is generally true whether we consider manual image interpretation procedures or computer assisted analyses. Processing sensor data into an interpretable format can be an effort entailing considerable thought, instrumentation, time, experience, and reference data.

6. The Multiple Data Users. Central to the successful application of any remote sensing is the person (or persons) using the remote sensor data from that system. The "data" generated by remote sensing procedures become "information" only if and when someone understands theirs generation, knows how to interpret them, and knows how best to use them. A through understanding of the problem at hand is paramount to the productive application of any remote sensing methodology. Also, no single combination of data acquisition and analysis procedures will satisfy the needs of all data users.

1.4 Electromagnetic Remote Sensing System

Remote sensing system can be thought as many respects, but the one that is currently being operated to assist in inventorying, mapping, and monitoring earth resources is electromagnetic remote sensing. Lillesand and Kiefer (1979) schematically illustrated the generalized process and elements involved in electromagnetic remote sensing of earth resource (Figure 1-2).

The two basic processes involved are data acquisition and data analysis. The elements of the data acquisition process are:

(a) Source of energy;

(b) Propagation of energy through the atmosphere;

(c) Energy interactions with earth surface features;

(d) Airborne and/or spaceborne sensors;

(e) Data Product. It is the generation of sensor data in pictorial and/or numerical form.

In short, we use sensors to record variations in the way earth surface features reflect and emit electromagnetic energy.

The elements of the data analysis process are:

(f) Interpretation;

(g) Information products;

(h) Users.

The data analysis process involves examining the data using various viewing and interpretation devices to analyze pictorial data, and/or a computer to analyze numerical sensor data. Reference data about the resources being studied (such as soils maps, crop statistics, or field-check data) are used when and where available to assist in the data analysis. With the aid of the reference data, the analyst extracts information (f) about the type, extent, location, and condition of the various resources over which the sensor data were collected. This information is then presented (g) generally in the form of maps, tables, and a written discussion or report. Typical information products are such things as land use maps and crop area statistics. Finally, the information is presented to users (h) who apply it to their decision-making process.

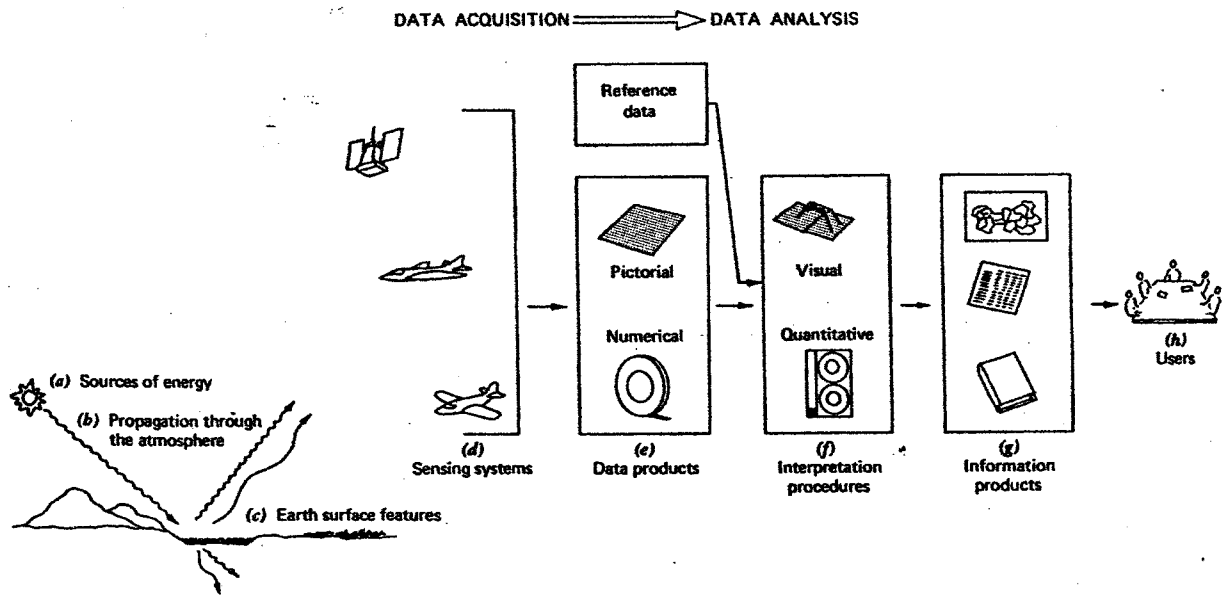


Figure 1-2: Electromagnetic remote sensing system of earth resource (From Lillesand and Kiefer, 1979)

Curran (1985) stated that a remote sensing system using electromagnetic radiation had four components: (1) a source, (2) interactions with the earth's surface, (3) interaction with atmosphere and (4) a sensor.

Source: the source of electromagnetic radiation may be natural like Sun's reflected light or the earth's emitted heat, or man-made, like microwave radar.

Earth's surface interaction: the amount and characteristics of radiation emitted or reflected from the earth's surface is dependent upon the characteristics of objects on the earth's surface.

Atmospheric interaction: electromagnetic energy passing through the atmosphere is distorted and scattered.

Sensor: the electromagnetic radiation that has interacted with surface of the earth and the atmosphere is recorded by a sensor, for example a radiometer or camera.

Landgrebe (2003) divided an Earth observational system into three basic parts: (1) the scene, (2) the sensor, and (3) the processing system as shown in Figure 1-3.

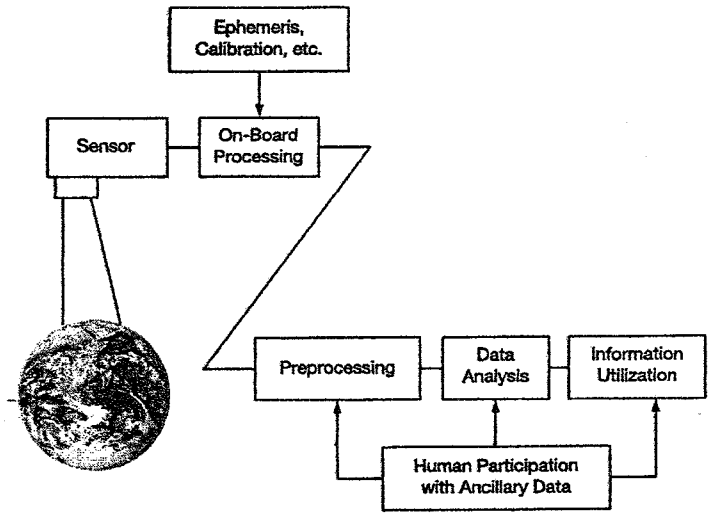


Figure 1-3: A conceptual view of an Earth observational system (From Landgrebe, 2003).

Scene: The scene consists of the Earth's surface and the intervening atmosphere. For systems in the optical portion of the spectrum that rely on solar illumination, the sun is included. This portion of the system is characterized by not being under human control, not during system design nor operation. However, its most defining characteristic is that it is by far the most complex and dynamic part of the system. It is easy to underestimate this complexity and dynamism.

Sensor: The sensor is the part of the system to collect the main body of data to be used. The sensor systems is usually characterized by being under human control during the system design phase, but less so or not at all during system operation.

Processing System: The processing system is that part of the system where data analyst is able to exercise the most control and have the most choice. One might divide the spectrum of possible degrees of human/machine participation in the analysis process into four parts: fully manual methods, machine aided manual methods, manually aided machine methods, and automatic analysis.

Jensen (2005) explained that the remote sensing data-collection and analysis procedures used for Earth resources application are often implemented in a systematic fashion referred as the remote sensing process. The procedures in the remote sensing process were here summarized and shown in Figure 1-4.

- The hypothesis to be tested is defined using a specific type of logic (e. g., inductive and deductive) and an appropriate processing model (e. g., deterministic, stochastic).
- In situ and collateral data necessary to calibrate the remote sensor data and/or judge its geometric, radiometric, and thematic characteristics are collected.
- Remote sensor data are collected passively or actively using analog or digital remote sensing instruments, ideally at the same time as the in situ data.
- In situ and remotely sensed data are processed using a) analog image processing, b) digital image processing, c) modeling and d) n-dimensional visualization.
- Metadata, processing lineage, and the accuracy of the information are the results communicated using images, graphs, statistical tables, GIS databases, Spatial Decision Support Systems (SDSS), etc.

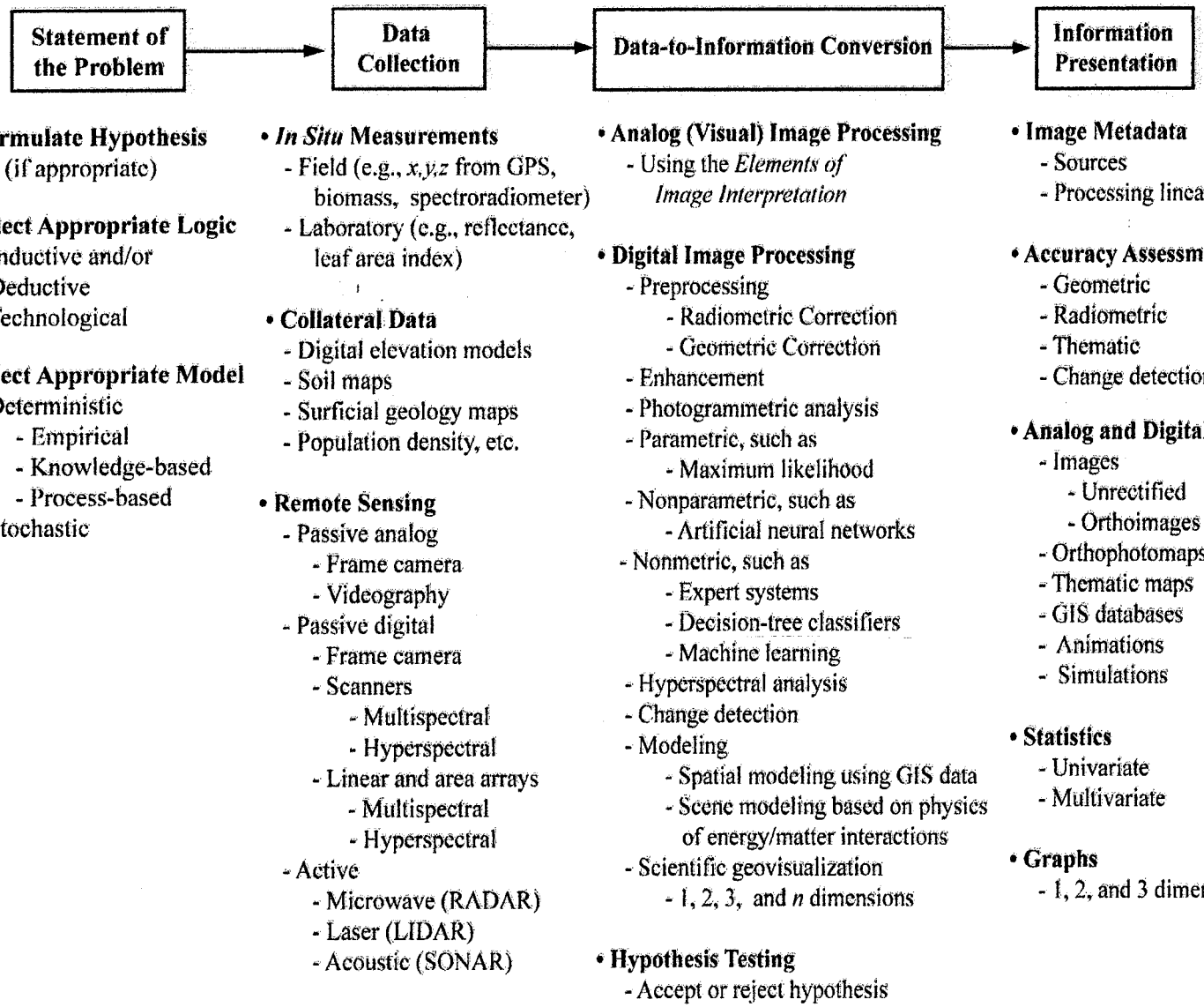


Figure 1-4: Remote sensing process when extracting information from remotely sensed data. (From Jensen, 2005)

1.5 Electromagnetic Radiation Principle

Remotely sensed images record the interaction of electromagnetic energy with the earth's surface. To develop proficiency in the interpretation of these images, it is necessary to understand the behavior of this energy as it reflected or emitted from the earth's surface through the atmosphere to the sensor. Therefore, knowledge of the properties and behavior of electromagnetic energy forms a foundation of remote sensing.

From basic physics, energy is the ability to do work (Figure 1-5). It is usually transferred from one point to another point by:

(a) Conduction. This involves atomic or molecular collisions.

(b) Convection. This is a corpuscular mode of transfer in which bodies of energetic material are themselves physically moved.

(c) Radiation. This is the only form in which electromagnetic energy may be transmitted either through a medium or a vacuum.

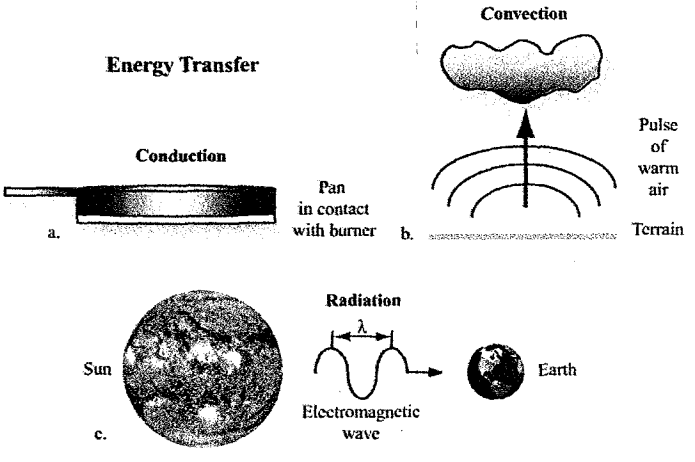


Figure 1-5: Types of energy transfer: conduction, convection and radiation (From Jensen, 2007)

In remote sensing we are primarily concerned with energy transfer by means of radiation. Electromagnetic radiation is one of the most useful force fields for remote sensing, forming a high-speed communications link between the sensor and remotely located substances (Suits, 1983).

In the case of electromagnetic radiation two models are necessarily to describe and elucidate its most characteristic: (a) the wave model and (b) the particular model (Lillesand and Kiefer, 1979; Barrett and Curtis; 1982 and Jensen, 2007).

1.5.1 The Wave Model. This typifies radiation through regular oscillary variations in the electric and magnetic fields surrounding a charged particle. Wave-like perturbations emanate from the source at the speed of light ($3x10^8$ m s^{-1}). They are generated by the oscillation of the particle itself. The two associated force-fields are mutually orthogonal, and both are perpendicular to the direction of advancement (Figure 1-6).

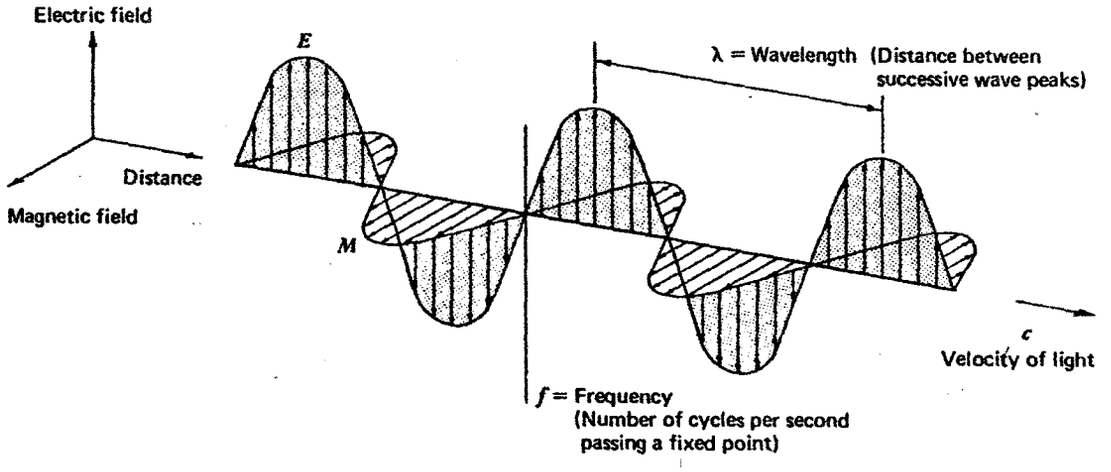


Figure 1-6: An electromagnetic wave. Components include a sinusoidal electro wave (E) and a similar magnetic wave (M) at right angles, both being perpendicular to the direction of propagation. (From Lillesand and Kiefer, 1979)

1.5.2 The Particular Model. This emphasizes aspect of the behavior of radiation which suggest that it is comprised of many discrete units. These are called "quanta" or "photons". These carry from the source some particle-like properties such as energy and momentum, but differ from all other particles in having zero mass at rest. It has been hypothesized, consequently that the photon is a kind of "basic particle".

In term of the wave model, waves obey the general equation from basic physics:

$$c = f\lambda \qquad (1\text{-}1)$$

where:

c is the speed of light, a constant (3×10^8 m s^{-1});

f is the wave frequency, cycle per second (Hertz, Hz);

λ is the wavelength (micrometers, μm)

Since c is constant, wave frequency is inversely proportional to wavelength and directly proportional to its speed of wave advancement.

In another term, particle model, the energy of a quantum is given as

$$E = h f \qquad (1\text{-}2)$$

where:

E is the energy of a quantum, Joule (J);

h is a Planck's constant, 6.226×10^{-34} J sec.

We can relate the wave and particle model of electromagnetic radiation by the common frequency value in both equations to obtain:

$$E = \frac{h \cdot c}{\lambda} \quad (1\text{-}3)$$

This tells us that the energy of a quantum is inverse proportional to its wavelength. The longer the wavelength involved, the lower its energy content. These relationships are fundamental to appreciation of the behavior of electromagnetic radiation.

In remote sensing, electromagnetic radiation, that occurs as a continuum of wavelengths and frequencies from short wavelength, high frequency to long wavelength, low frequency, is common categorized as electromagnetic spectrum. The boundaries between them are expressed in several different ways. Figure 1-7 is the one that be proposed by Suits (1983). It shows the extent of the electromagnetic spectrum, the various named bands, the transmittances of the earth's atmosphere to the electromagnetic radiation, and the effects caused by its interaction or presence.

The sun is the most obvious source of electromagnetic radiation for remote sensing. However, all matter at temperatures above absolute zero (0°K or -273°C) continuously emits electromagnetic radiation. Thus, terrestrial objects are also sources of radiation though it is considerably different magnitude and spectral composition than the sun (Lillesand and Kiefer, 1979).

Barrett and Curtis (1982) stated that a useful concept widely used by the physicists in radiation studies, is that of blackbody, a model (perfect) absorber and radiator of electromagnetic radiation. A blackbody is conceived to be an object or substance which absorbs all the radiation incident upon it, and emits the maximum amount of radiation at all temperatures. Although there is no known substance in the natural world with such a performance, the blackbody concept is invaluable for the formulation of laws by comparison with which the behavior of actual radiators may be assessed. They also summarized the principal laws that relate to this concept as following:

1.5.3 Stefan-Boltzmann' Law. This states that the total emissive power of a blackbody is proportional to the fourth power of its absolute temperature (T). This can be expressed as:

$$M = \sigma T^4 \quad (1\text{-}4)$$

where:

M is the radiant exitance in W m^{-2}

σ is the Stefan-Boltzmann Constant, 5.6697×10^{-8} W m^{-2} °K^{-4}

T is absolute temperature (°K) of emitting material.

This relationship applied to all wavelengths shorter than the microwave. In the microwave region radiant exitance varies as a direct function of T (°K). The Stefan-Boltzmann's Law is that hot radiators emit more energy per unit area than cooler ones.

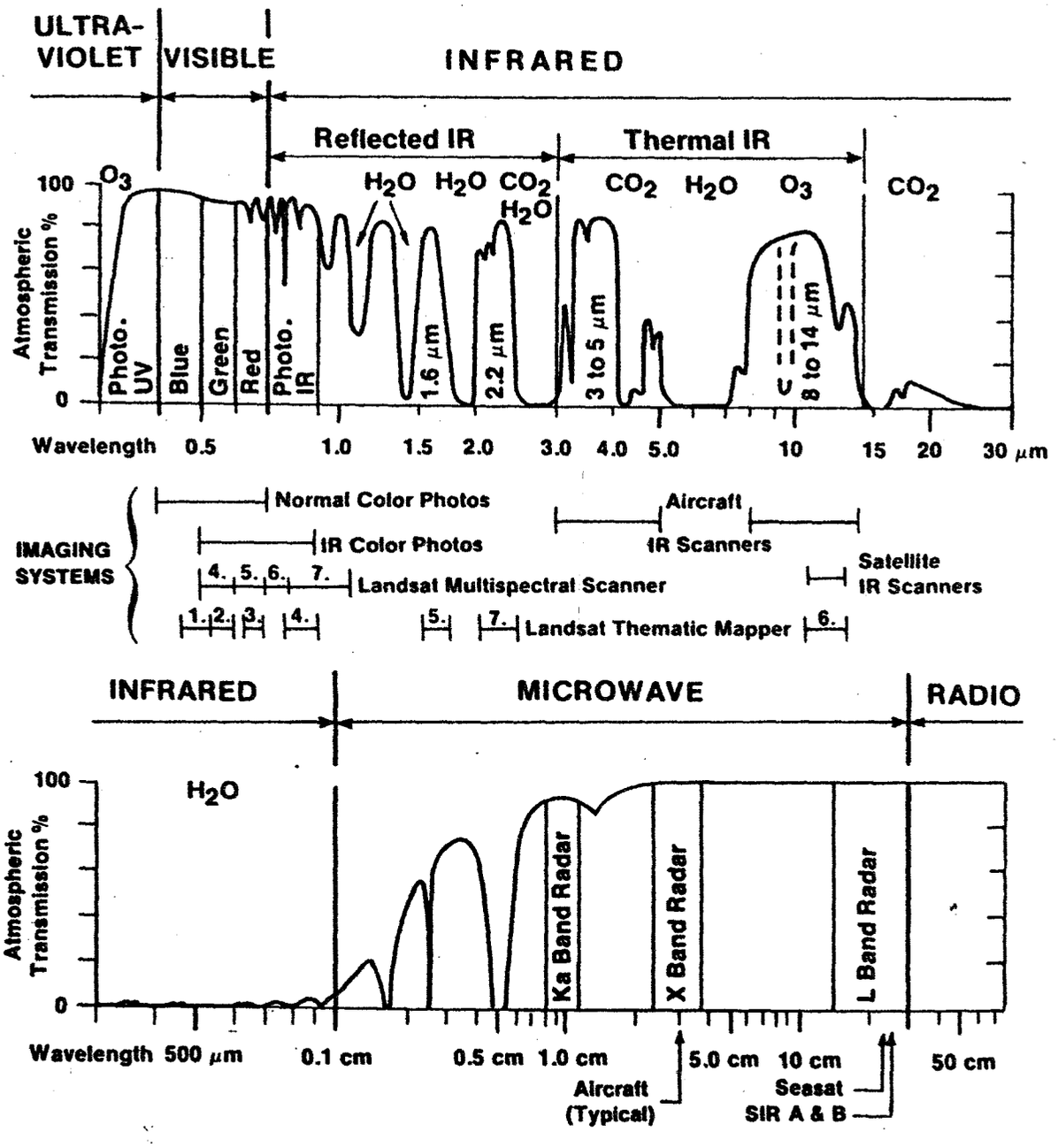


Figure 1-7: Electromagnetic spectrum diagrams of the visible and infrared regions (upper) and the microwave region (lower) showing atmospheric windows. Wavelength bands of commonly used remote sensing systems are indicated. Gases responsible for atmospheric absorption are shown (Form Suits, 1983).

1.5.4 Kirchhof's Law. Since no real body is a perfect emitter, its exitance is less than that of a blackbody. Clearly it is often useful to known how the real exitance (W) of a radiator compares with that which would be anticipated from a corresponding perfect radiator (W_b). This may be established by evaluating the ratio W/W_b, which gives the emissivity (ε) of the real body. Thus, for the general case, it was said that:

$$W = \varepsilon W_b \qquad (1\text{-}5)$$

The emissivity of a real blackbody would be 1, whist the emissivity of a body absorbing none of the radiation upon it would be 0. Between these two limiting values, the grayness of real radiators can be assessed, frequently to two decimal places.

1.5.5 Wien's Displacement Law. This states that the wavelength of peak radiant exitant (λ_{max}) of a blackbody is inversely proportional to its absolute temperature (T):

$$\lambda_{max} = C_3 / T \qquad (1\text{-}6)$$

where:

C_3 is a constant equal to 2897 μmK;

T is absolute temperature (°K).

This equation tells us that, as the temperature of a blackbody increase, so that dominant wavelength of emitted radiation shifts towards the short wavelength and of the spectrum. This can be exemplified by reference to the Sun and Earth as radiating bodies. For the Sun, with a mean surface temperature of about 6000 °K, λ_{max} = 0.5 μm. For the Earth, with a surface temperature of about 300 °K on a warm day, λ_{max} = 9.0 μm (See also Figure 1-8).

1.5.6 Planck's Law. This more complex statement describes accurately the spectral relationships between the temperature and radiative properties of a blackbody. In one form, Planck's Law may be expressed by

$$M_\lambda = \frac{\varepsilon \cdot c_1}{\lambda^5 \left(e^{c_2 / \lambda T} - 1 \right)} \qquad (1\text{-}7)$$

where:

M_λ is spectral radiant exitance

ε is emittance (emissivity)

C1 is first radiation constant, 3.74151×10^8 W-m^{-2}-μm^4

C2 is second radiation constant, 1.433879×10^4 μm-K

T is absolute temperature (°K).

A useful feature of Planck's Law is that it enables us to assess the proportions of the total radiant existence, remote sensor design, and interpretation of remote sensing observations.

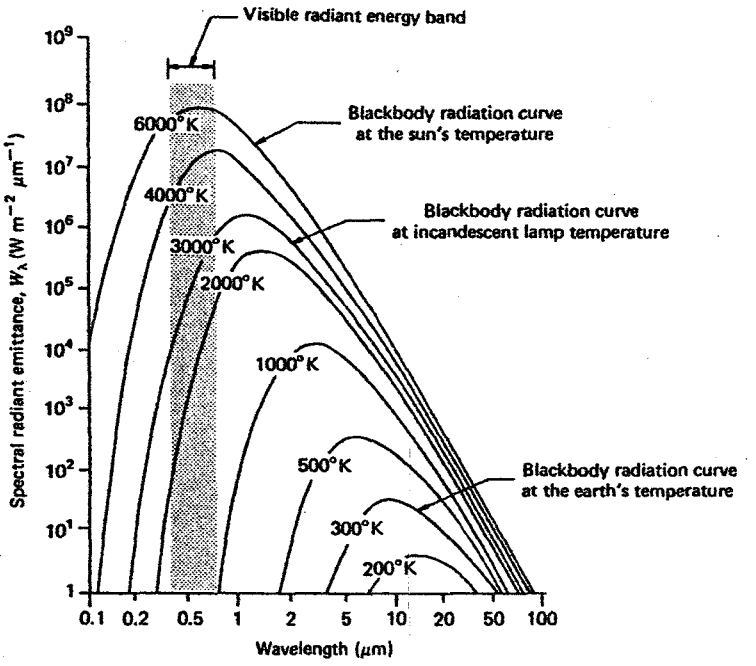


Figure 1-8: Spectral distribution of energy radiated from blackbody of various temperatures (From Lillesand and Kiefer, 1979).

The interrelationships between temperature, wavelength, frequency, radiant energy, radian exitance and the point of maximum spectral radiant exitance can be summarized by Figure 1-9.

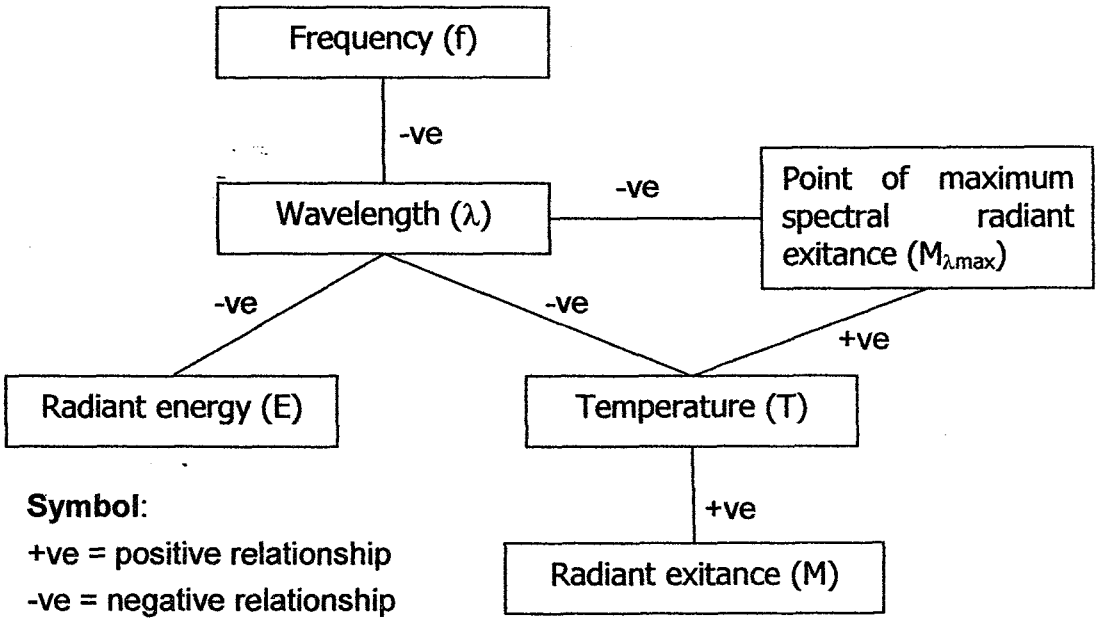


Figure 1-9: Interrelationship between temperature, wavelength, frequency radiant energy, radiar exitance and the point of maximum spectral radiant exitance (From Curran, 1985).

1.6 Energy Interactions in the Atmosphere

The earth's atmosphere that intervenes between the sensor and its target plays important role in remote sensing system. In general, atmospheric effects may have substantial impact upon the quality of images and data that the sensors generate. Therefore, the practice of remote sensing requires knowledge of interactions of radiations with the atmosphere.

Three physical processes in the atmosphere are considered in remote sensing including: scattering, absorption and refraction.

1.6.1 Scattering

Scattering is the redirection of electromagnetic energy by particles suspended in the atmosphere or by the large molecules of atmospheric gases (Campbell, 1987).

Barrett and Curtis (1982) stated that although the speed of electromagnetic radiation is un-effect by the atmosphere, this atmosphere may affect several of the other characteristics of this form of energy propagation. These include:

a) The direction of radiation;

b) The intensity of radiation;

c) The wavelength and frequency of the radiation received by a target, at the base of the atmosphere;

d) The spectral distribution of this radiant energy.

The attenuation which results from scattering by particles suspended in the atmosphere is related to the wavelength of radiation, the concentration and diameters of the particles, the optical density of the atmosphere and its absorptivity. The common types of scatter are Rayleigh, Mie and Nonselective scattering as shown in Figure 1-10.

(a) Rayleigh scattering. This mostly involves molecules and other tiny particles with diameters much less than the radiation wavelength in question. It is characterized by an inverse fourth power dependence on wavelength. For example, ultraviolet radiation (about one-quarter the wavelength of red light) is scattered sixteen times as much. This helps to explain the dominance of arrange and red at sunset when the sun is low in the sky: the shorter wavebands of visible light are cut out by a combination of atmospheric absorption and powerful scattering.

Landgrebe (2003) state some of the factors influencing the scattering process can be seen from the equation for the Volume Scattering Coefficient (σ_λ) indicating the relative amount of the radiation that is scattered, as follow:

$$\sigma_{\lambda} = 4\pi^2 \frac{NV^2\left(n^2 - n_0{}^2\right)^2}{\lambda^4\left(n^2 - n_0{}^2\right)^2} \tau_v(\lambda) \qquad (1\text{-}8)$$

where:

N is number of particle/cm3

V is volume of scattering particles

λ is radiation wavelength

n is refractive index of particles

n_0 is refractive index of medium.

(b) Mie scattering. This occurs when the atmosphere contains essentially spherical particles whose diameters approximates to the wavelengths of radiation in question. Water vapor and particles of dust are the main agents which scatter visible light.

(c) Nonselective scattering. Here particles with diameters several times the radiation wavelengths are involved. Water droplets, for example, with diameters ranging commonly from 5-100 μm scatter all wavelengths of visible light (0.4-0.7 μm) with equal efficiency. As a consequence clouds and fog appear whitish, for a mixture of all colors in approximately equal quantities produce white light.

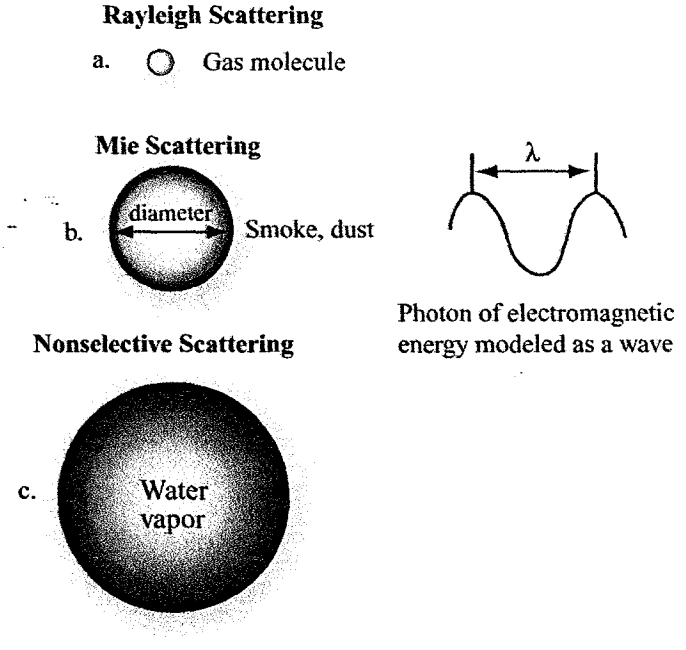


Figure 1-10: Atmospheric scattering (From Jensen, 2007)

1.6.2 Absorption

Absorption is the retention of radiant energy by a substance or a body. In the real world it involves the transformation of some the incident radiation into heat, and the subsequent re-emission of that energy at a longer wavelength (Barrett and Curtis, 1982)

Absorption of radiation occurs when the atmosphere prevents or strongly attenuates, transmission of radiation or its energy through the atmosphere. Three gases responsible for most absorption of solar radiation: ozone, carbon dioxide, and water vapor (Lillesand and Kiefer, 1979; Barrett and Curtis, 1982; Campbell, 1987).

(a) Ozone (O_3) is formed by the interaction of high-energy ultraviolet radiation with oxygen molecules (O_2) high in the atmosphere (maximum concentrations of ozone are quite low, ozone plays an important role in the earth's energy balance. Absorption of the high-energy, short wavelength, portions of the ultraviolet spectrum prevents transmission of the radiation to the lower atmosphere.

(b) Carbon dioxide (CO_2) occurs in low concentrations (about 0.03 % by volume of a dry atmosphere), mainly in the lower atmosphere. Aside from local variation caused by volcanic eruptions and human activities, the distribution of CO_2 in the lower atmosphere is probably relative uniform. Carbon dioxide is important in remote sensing because of effectively absorbs radiation in the mid and far infrared regions of the spectrum. Its strongest absorption occurs in the regions from about 13 to 17.5 μm, in the far infrared.

(c) Water vapor (H_2O) is commonly present in the lower atmosphere in amounts that from 0 to 3 % by volume. The abundance of water vapor varies greatly from time to time and from place to place. Consequently, the role of atmospheric water vapor, unlike those of ozone and carbon dioxide, changes greatly with time and location. It may be almost insignificant in a desert or a dry air mass but highly significant in humid climates and in moist air masses. Furthermore, water vapor is several times as effective in absorbing radiation as are all other atmospheric gases combined. Two of most important regions of absorption are in several bands between 5.5 and 7.0 μm, and above 27.0 μm; absorption in these regions can exceed 80 % if the atmosphere contains appreciate amounts of water vapor. Figure 1-11 shows atmospheric absorption by various gases.

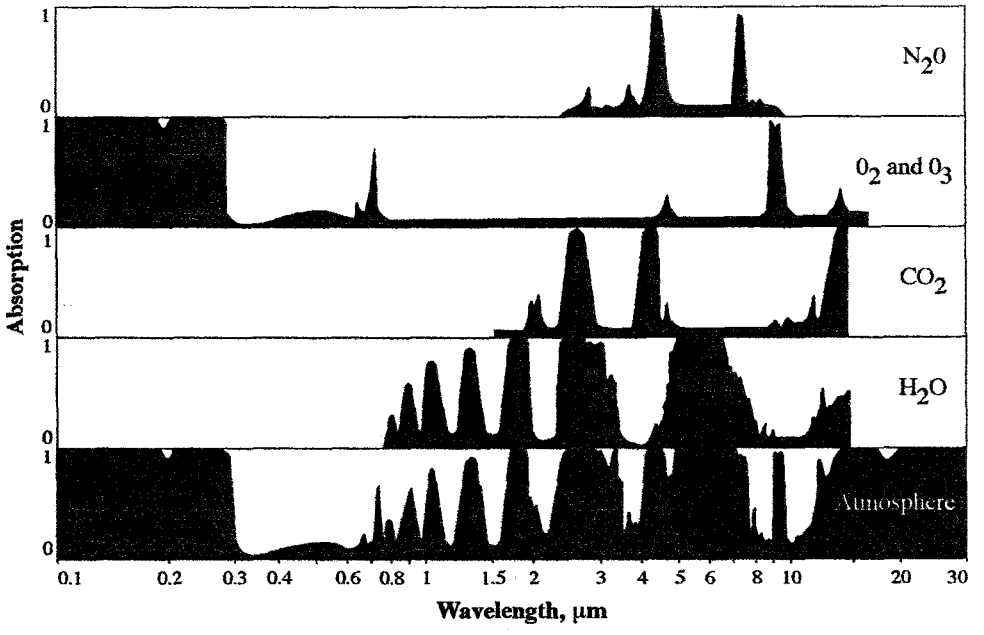


Figure 1-11: Atmospheric absorption by various gases in the region from 0.1 to 30 μm (From Jensen, 2007).

1.6.3 Refraction

Campbell (1987) defined refraction is the bending of light rays at the contact between two media that transmit light. Therefore, when electromagnetic radiation passed from one medium to another, refraction occurs in response to the contrasting densities of the media (Figure 1-12). A measure of this is given by the Index of Refraction (n) (Barrett and Curtis, 1982).

The index of refraction (n) is defined as the ratio between the velocity of light in a vacuum (c) to its velocity in the medium (c_n):

$$n = c/c_n \qquad (1\text{-}9)$$

In a non-turbulent atmosphere (which can be conceived as a series of layers of gasses each of a different density) the refraction of radiation is predictable. For as Snell's law states, there is a constant relationship for a given frequency of light between the index of refraction and the sine of the angle between the ray and the interface between each pair of adjacent layers, where:

$$n \sin \emptyset_1 = n' \sin \emptyset_2 \qquad (1\text{-}10)$$

where n and n' are the indices of refraction of the first and second media, respectively, and $\emptyset_1$ and $\emptyset_2$ are angles measured with respect to the surface normal (See Figure 1-12).

However, problems do arise, where a turbulent atmosphere is involved. Turbulent motions are essentially random, and their effects upon radiation are consequently unpredictable (Barrett ar Curtis, 1982).

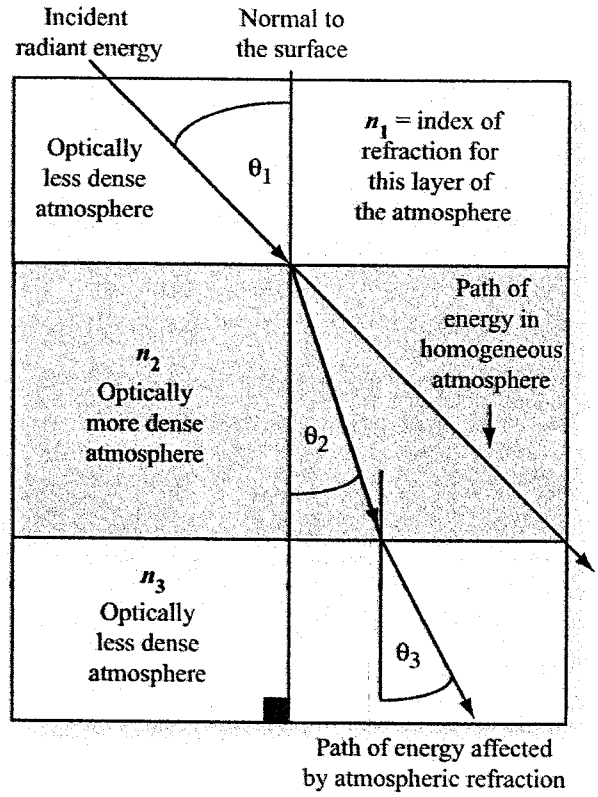


Figure 1-12: Atmospheric refraction (From Jensen, 2007)

1.7 Atmospheric Windows

Atmospheric absorption results in the effective loss of radiant energy to atmospheric constitutes. The atmosphere selectively transmits energy of a certain wavelengths. The wavelength ranges in which the atmosphere is particularly transmissive of energy are referred as atmospheric windows (Lillesand and Kiefer, 1979).

Figure 1-13 shows the interrelationship between energy sources and atmospheric absorption characteristics. Figure 1-13(a) shows the spectral distribution of the energy emitted by the sun and by the earth features. These two curves represent the most common sources of energy used in remote sensing. In Figure 1-13(b), spectral regions in which the atmosphere blocks energy are shaded. Remote sensing data acquisition is limited to the non-blocked spectral regions, called atmospheric windows. In Figure 1-13(c), the spectral sensitivity range of the eye coincides both with an atmospheric window and the peak level of energy from the sun. Emitted heat energy from the earth is sensed through the window at 3 to 5 μm and 8 to 14 μm using such devices as thermal scanners. Multispectral scanners sense simultaneously through multiple, narrow wavelength ranges that can be located at various points in the visible through the thermal spectral region. Radar and passive microwave systems operates through a window in the 1 mm to 1 m region.

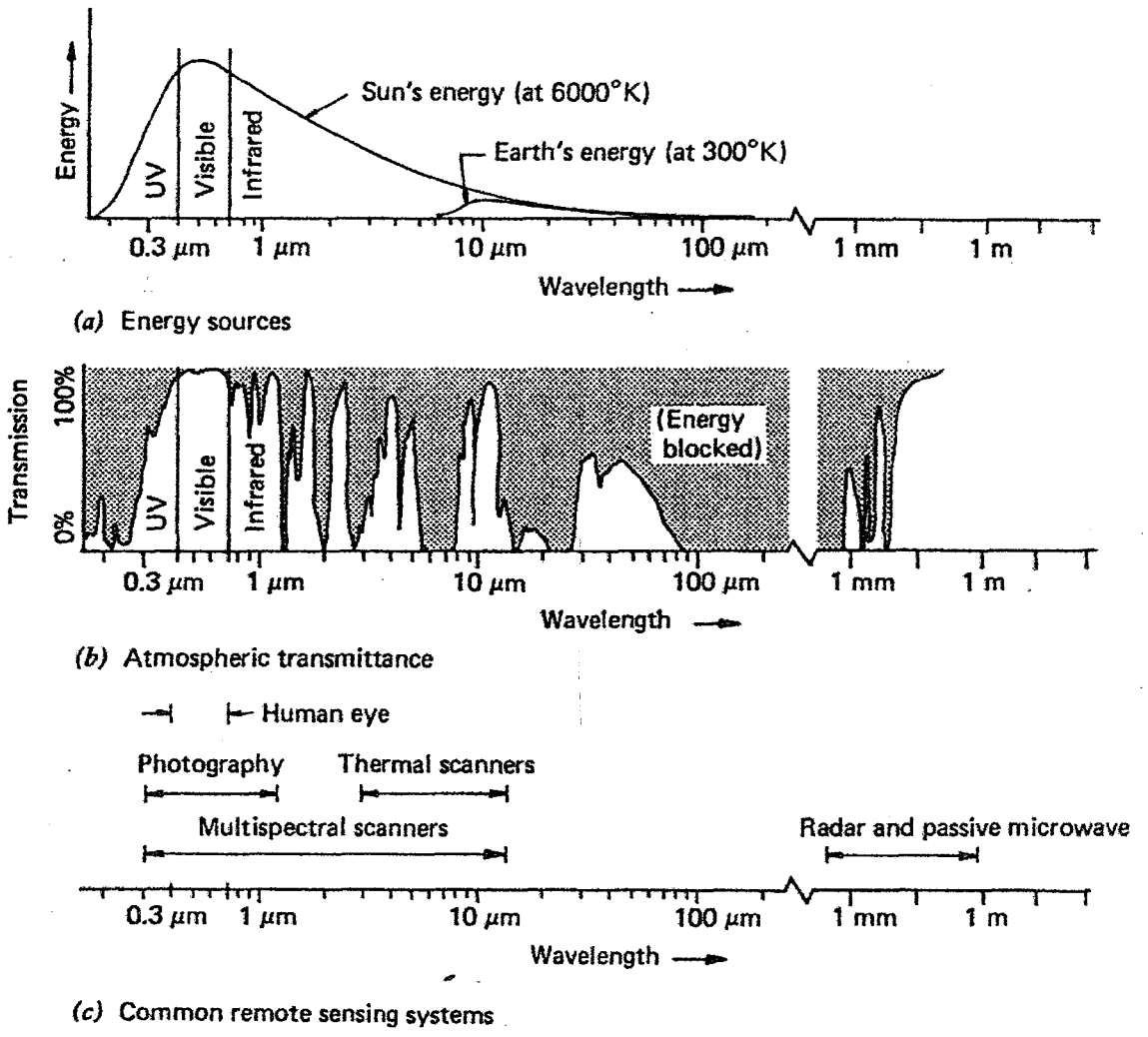


Figure 1-13: Spectral characteristics of energy sources, atmospheric effects, and sensing system (Note that wavelength scale logarithmic) (From Lillesand and Kiefer, 1979)

1.8 Interactions with the Earth Surface Feature

After the radiance energy penetrates through the atmosphere, it reaches the earth surface feature. Three fundamental energy interactions with the feature are possible. It must be reflected, absorbed and/or transmitted.

Lillesand and Kiefer (1979) stated that when electromagnetic energy is incident on any given earth surface feature; three fundamental energy interactions with the feature are possible. Figure 1-14 illustrated this principle for an element of the volume of water body. They also stated that when applying the principle of conservation of energy, the interrelationship between three energy interactions as:

$$E_I(\lambda) = E_R(\lambda) + E_A(\lambda) + E_T(\lambda) \qquad (1\text{-}11)$$

where E_I denotes the incident energy, E_R denotes the reflected energy, E_A denoted the absorbed energy and E_T denoted the transmitted energy. This is an energy balance equation expressing the interrelationship between the mechanisms of reflection, absorption, and transmission.

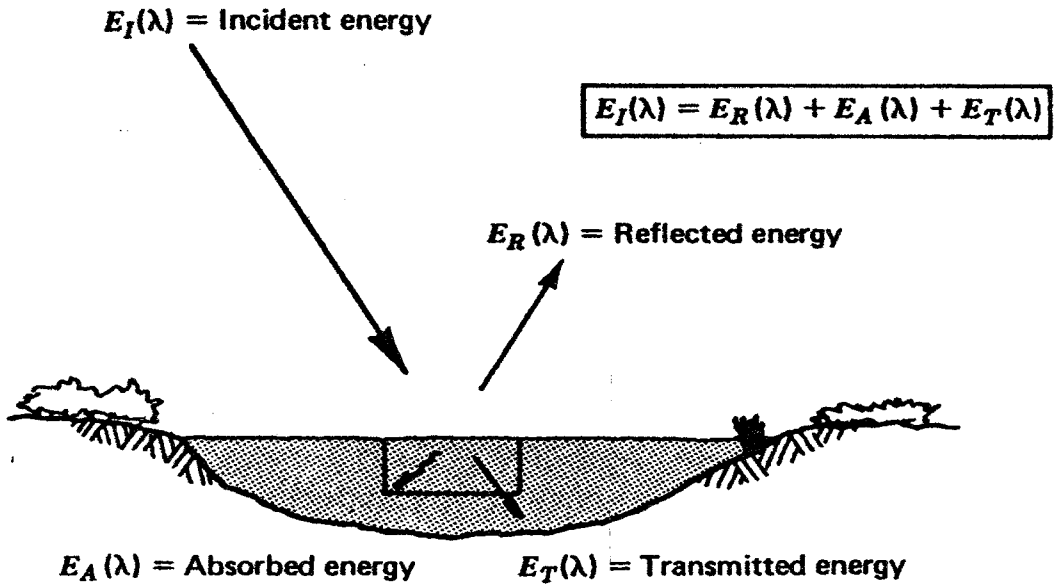


Figure 1-14: Basic interactions between electromagnetic energy and the earth surface feature (From Lillesand and Kiefer, 1979)

The properties accounted for by each mechanism depends upon the nature of the surface, the wavelength of the energy, and the angle of illumination (Campbell, 1987)

Because many remote sensing systems operate in the wavelength regions in which reflected energy predominates, the reflectance properties of earth features are very important. Thus the energy balance is often expressed in the following form

$$E_R(\lambda) = E_I(\lambda) - [E_A(\lambda) + E_T(\lambda)] \quad (1\text{-}12)$$

In remote sensing, the reflectance characteristics of the earth surface may also be quantified by measure as a function of wavelength, and is called *spectral reflectance* (Lillesand and Kiefer, 1979). Spectral reflectance (R_λ) in percent is mathematically defined as

$$R_\lambda = \frac{E_R(\lambda)}{E_R((\lambda)} = \frac{Energy\ of\ wavelength\ a\ reflected\ from\ the\ object}{Energy\ of\ wavelength\ a\ reflected\ from\ the\ object} \times 100 \quad (1\text{-}13)$$

A graph of the spectral reflectance of an object as a function of wavelength is termed a *spectral reflectance curve* (Figure 1-15). Theoretically, each object on the earth's surface has its own unique spectral or radiation reflectance characteristics. Thus, we can identify the features of earth's surface on the basis of their spectral properties.

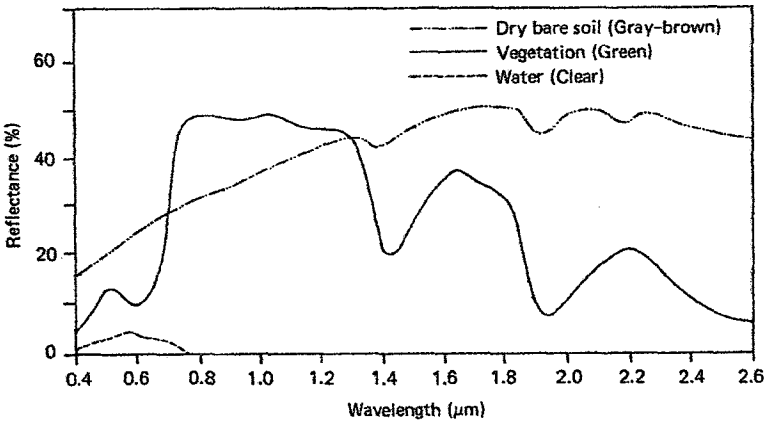


Figure 1-15: Typical spectral reflectance curve for vegetation, soil and water (From Lillesand et al., 2004)

Curran (1985) stated that whether a remote sensing can detect such variation in reflectance between objects is dependent upon four interrelated factors: (1) the radiometric resolution of the sensor, (2) the amount of atmospheric scatter, (3) the surface roughness of the objects, and (4) the spatial variability of reflectance with in the scene.

First, sensors vary in their ability to detect difference in radiance. For example, on board the satellite Landsat 4 the scanning radiometer called Thematic Mapper (TM) can detect 256 levels of radiance and the scanning radiometer called the Multispectral Scanner (MSS) can detect 64 levels of radiance.

Second, atmospheric scatter increase the amount of radiance received by the sensor for each object and as a result the contrast between object is reduced.

Third, surface roughness is of importance because the surface needs to be rough enough to allow radiation to interact with the surface of objects. If the surface of the object is smooth and radiation is reflected without interaction, then information will be transmitted to the sensor. Jensen (2007) stated that there are various types of reflecting surfaces. When *specular reflection* occurs, the surface from which the radiation is reflected is essentially smooth. If the surface is rough, the reflected rays go in many directions, depending on the orientation of the smaller reflecting surfaces. This diffuse reflection does not yield a mirror image, but instead produces *diffused radiation*. White paper, white powders and other materials reflect visible light in this diffuse manner. If the surface is so rough that there are no individual reflecting surfaces, then scattering may occur. Lambert defined a *perfectly diffuse surface*, hence the commonly designated *Lambertian surface* is one for which the radiant flux leaving the surface is constant for any angle of reflectance to the surface normal (Figure 1-16).

Fourth, the spatial variability of the scene is of important because every sensor has what is termed "a point spread function". This means that the radiance recorded from an area of ground also contains radiance from the surrounding areas.

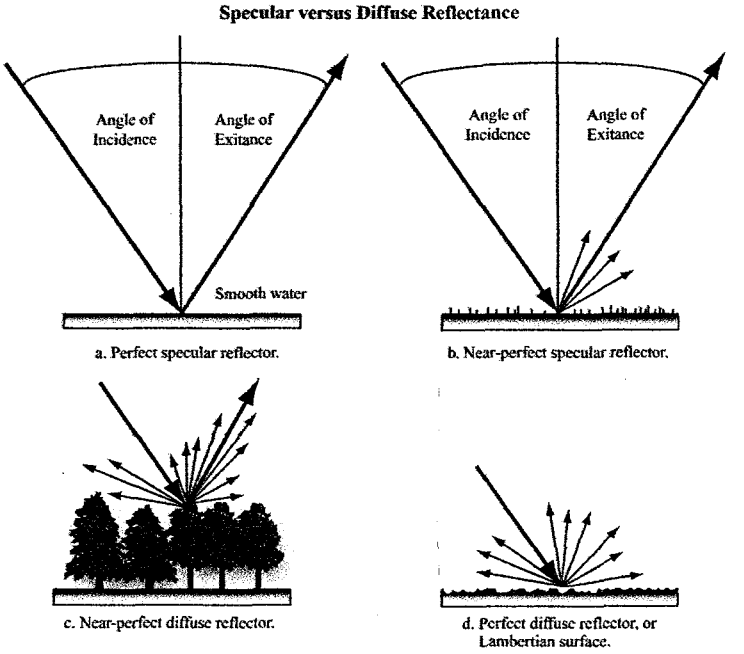


Figure 1-16: The nature of specular and diffuse reflectance (From Jensen, 2007)

Both the reflected and incident radiation at a surface can be measured in three different ways include (1) Bi-directional reflectance (2) Directional reflectance and (3) Hemisphere reflectance (McCloy, 2006) as shown in Figure 1-17. Characteristics of three different types are summarized based on McCloy (2006) as follows:

A. Bi-directional reflectance. If the observer is looking at a surface that is illuminates by torchlight then the source of energy comes from one direction, and the surface is being observed at another direction. The ratio of these two energies is called bi-directional reflectance factor of the surface, being dependent on the two directions involved. Illumination of surfaces by the sun, involving predominantly illumination from the one direction, and observation at a sensor is a close approximation to bi-directional reflectance. It is an approximation because the definition of bi-directional reflectance assumes a negligible FOV for both the incident and reflected radiation, and because skylight contributes to illuminations of the target. For non-Lambertian surfaces, the bi-directional reflectance may change with changes in the sun-surface-sensor geometry. In consequence the reflectance of some surfaces may change across an image and form image to image. Most forms of active sensors, or sensors that have their own energy source to illuminate the target, use a point source and sense in one direction so that they represent perfect bi-

directional reflectance. The dependent of the bi-directional reflectance on the incident and reflected energy and reflected angle is shown as:

$$\text{Bi-directional reflectance factor} = \frac{L(\Omega)}{\pi \times I(\Omega_o)} \quad (1\text{-}14)$$

where:

L is the radiant of reflected energy in direction Ω

I is the irradiance of the incident energy from direction Ω_0.

The bi-directional reflectance factor will only be the same for all combinations of azimuth and elevation for both incident and reflectance radiation, if the surface is a perfect Lambertian reflector. Thus has given rise to the concept of bi-directional reflectance distribution function (BRDF). The BRDF is a map of the bi-directional reflectance factor for all combination of altitude and azimuth of both the incident and reflected radiation. Such a map is difficult to portray graphically. As a consequence, only the portion of the BRDF is often portrayed, usually the bi-directional reflectance factor with variations in the reflected or sensed angle, at constant incidence angle and at one waveband. The BRDF of three surfaces are shown in Figure 1-18 to 1-20.

The BDRF shows how a surface varies from the Lambertian ideal, how the surface varies from other surface, and indeed how the surface can change over time. With vegetation, the BRDF can also vary during the day and over longer time periods, since the canopy structure does change over time, and can change during the day (Figure 1-21). Analysis of variations in BRDF over time, both within a cover type, and between cover types, can indicate strategies for the acquisition of data so as to maximize the difference in reflectance between selected surface types.

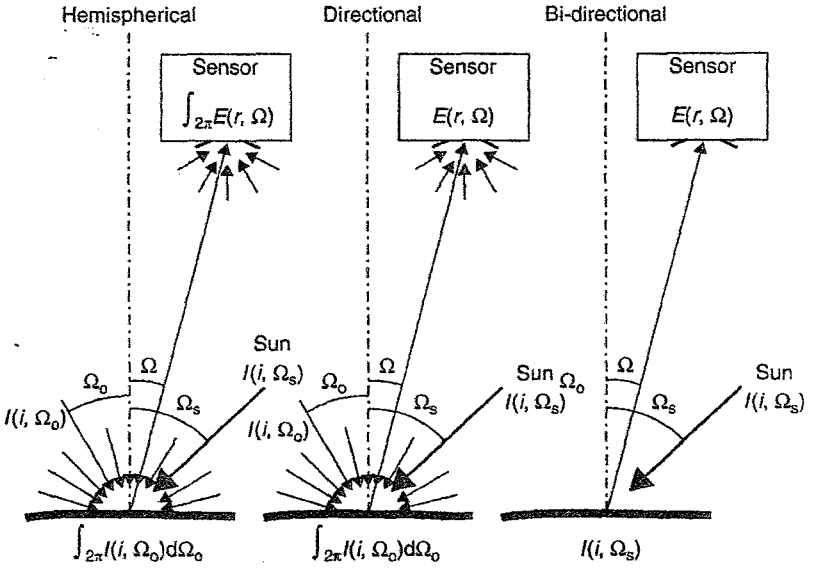


Figure 1-17: Hemispherical, directional and bi-directional reflectance (From McCloy, 2006)

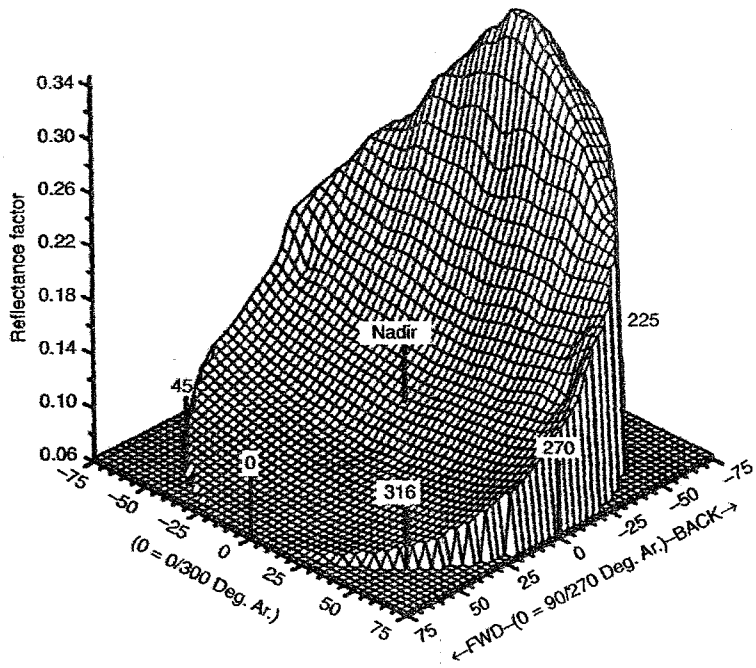


Figure 1-18: The bi-directional reflectance function (BRDF) of a recent ploughed bare field at 25° solar elevation for 662 nm wavelength. (From McCloy, 2006)

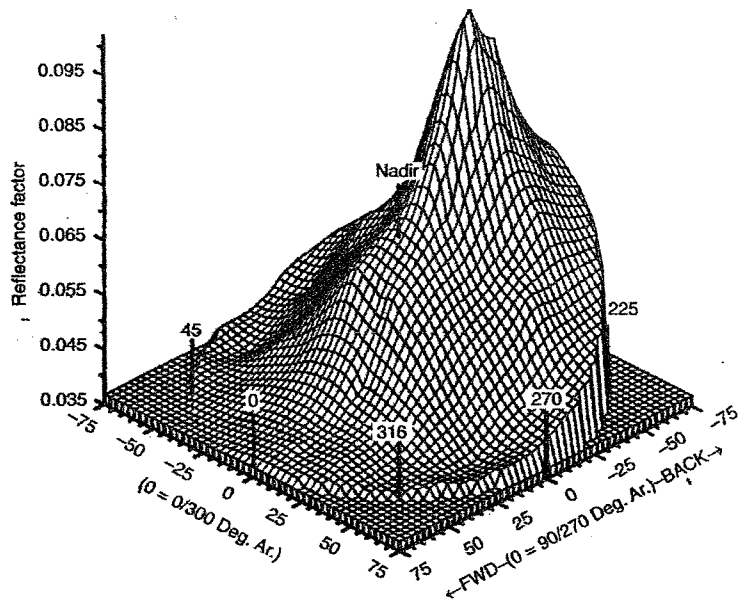


Figure 1-19: The bi-directional reflectance function (BRDF) of a sand shinnery oak rangelands community at 31° solar zenith angle at = 662 nm. (From McCloy, 2006)

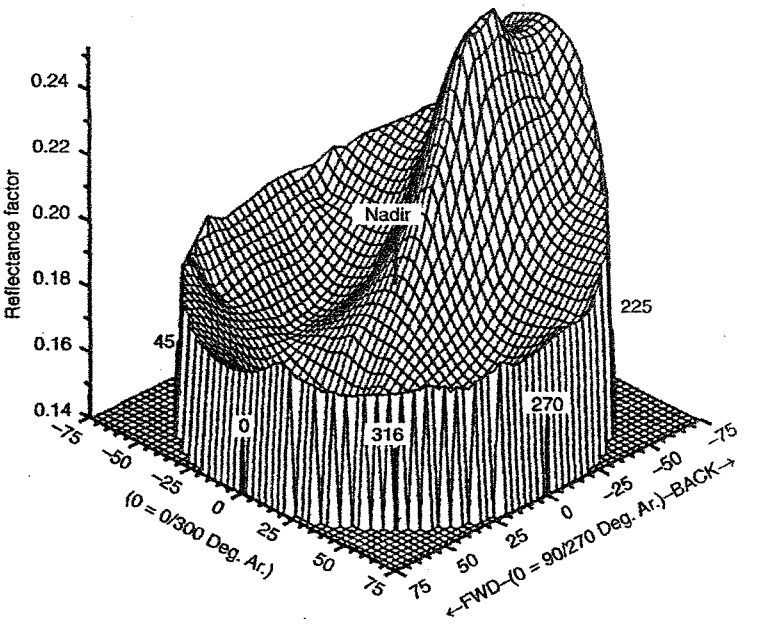


Figure 1-20: The bi-directional reflectance function (BRDF) of a sand shinnery oak rangelands community at 31° solar zenith angle at = 862 nm. (From McCloy, 2006)

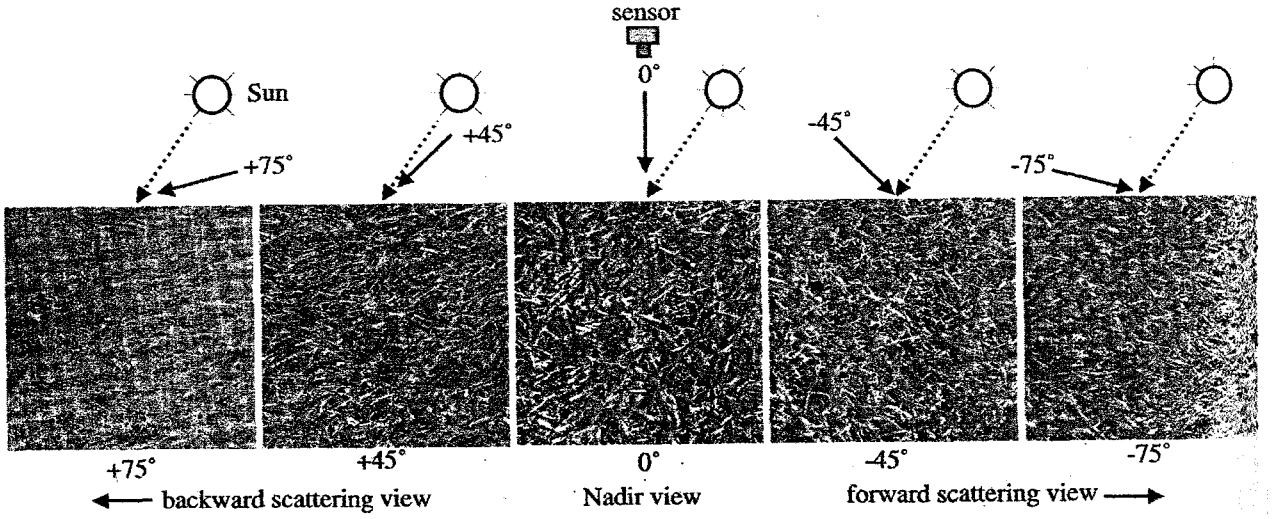


Figure 1-21: The bi-directional reflectance effects (From Jensen, 2007)

B. Directional reflectance. If the observer views the surface under overcast conditions when the main source of illumination is scattered skylight, then the illumination is approximately equal from all directions, but sensor is in one direction relative to the surface. Reflectance of this form is called directional reflectance. The directional reflectance is the ratio of the single directional reflected energy to the total irradiance incident on the surface giving the hemispherical directional reflectance facto as:

$$\text{Hemispherical directional reflectance factor} = \frac{L(\Omega)}{\pi \times \int_{2\pi} I(\Omega_o) d\Omega_o} \qquad (1\text{-}15)$$

C. Hemispherical reflectance. Hemispherical reflectance occurs when both the incident and reflected energy is measured over the whole hemisphere. Hemispherical reflectance is of little concern in relation to sensor, but is of great interest in the modeling of light interactions within the surface and in assessing the impact of reflectance on aerial photography and other imagery that is taken with sensors that have a wide FOV. Hemispherical reflectance should be identical to bi-directional reflectance for Lambertian surfaces, and provides an average reflectance value for specular surfaces. Comparison of hemispherical and bi-direction reflectance values can thus indicate how close surface are to the Lambertian model. For canopy modeling and inversion, hemispherical reflectance is defined as the ratio of total (hemispherical) reflected energy from a location to the total (hemispherical) incident energy. This bi- hemispherical reflectance factor is given by:

$$\text{Bi-Hemispherical directional reflectance factor} = \frac{\int_{2\pi} L(r,\Omega)d\Omega}{\int_{2\pi} I(\Omega_o)d\Omega_o} \qquad (1\text{-}16)$$

In conclusion, Jensen (2007) stated that the amount of electromagnetic radiance, L (watts m-2 sr-1; watts per meter squared per steradian) recorded within the IFOV of an optical remote sensing system (e.g., a picture element in a digital image) is a function of:

$$L = f\left(\lambda, s_{x,y,z}, t, \theta, P, \Omega\right) \qquad (1\text{-}17)$$

where:

- λ is wavelength (spectral response measured in various bands or at specific frequencies). Wavelength (λ) and frequency (ν) may be used interchangeably based on their relationship with the speed of light (c): where $c = \lambda \times \nu$.
- $s_{x,y,z}$ is x,y,z location of the picture element and its size (x, y)
- t is temporal information, i.e., when and how often the information was acquired
- θ is set of angles that describe the geometric relationships among the radiation source (e.g., the Sun), the terrain target of interest (e.g., a corn field), and the remote sensing system
- P is polarization of back-scattered energy recorded by the sensor
- Ω is radiometric resolution (precision) at which the data (e.g., reflected, emitted, or back-scattered radiation) are recorded by the remote sensing system.

Lillesand et al. (2004) illustrated the basic components of radiation entered through the sensor when records reflected solar energy as shown in Figure 1-22. This figure provides an initial frame of reference for understanding the nature of atmospheric affecting to the spectral response

pattern. They stated that the atmosphere affects the "brightness," or radiance, recorded over given point on the ground in the two almost contradictory ways. First, it attenuating a ground object and being reflected from the object. Second, the atmosphere acts as a reflector itself, adding a scattered, extraneous *path radiance* to the signal detected by the sensor. By expressing these two atmospheric effects mathematically, the total radiance recorded by the sensor may be related to the reflectance of the ground object and the incoming radiation or irradiance using the equation:

$$L_{tot} = \frac{\rho ET}{\pi} + L_p \qquad (1\text{-}18)$$

where:

L_{tot} is total spectral radiance measured by sensor,

P is reflectance of object,

E is irradiance an object, incoming energy,

T is transmission of atmosphere

L_p is path radiance, from the atmosphere and not from the object.

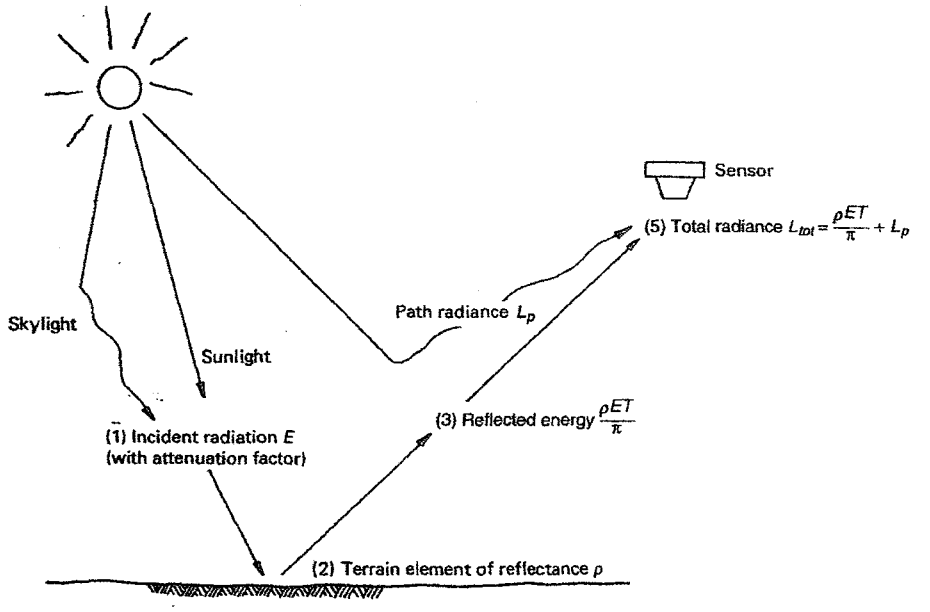


Figure 1-22: Atmospheric effects influencing the measurement of reflected solar energy. (From Lillesand et al, 2004)

It should be note that all of the above factors depend on wavelength. Also, the irradiance (E) stems from two sources: (1) directly reflected "sunlight: and (2) diffuse "skylight" which is sunlight that has been previously scattered by the atmosphere. The relative dominance of sunlight versus skylight in any given images is strongly dependent on whether condition (e.g., sunny vs.

hazy vs. cloudy). Likewise, irradiance varies with the seasonal changes in solar elevation angle (Figure 1-23) and the change distance between the Earth and Sun.

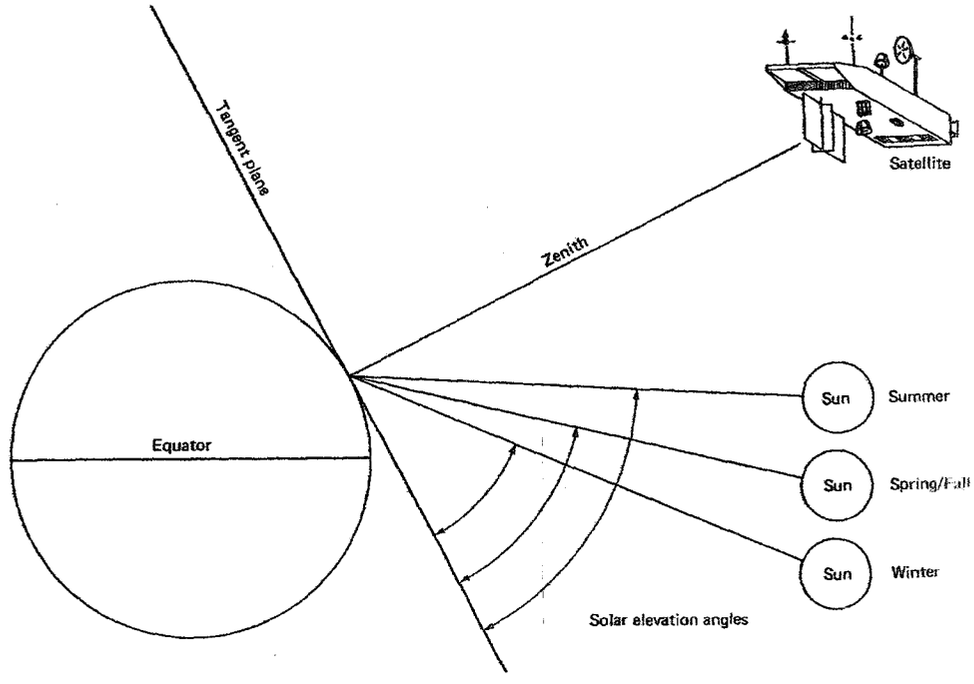


Figure 1-23: Effects of seasonal change on solar elevation angle. (The solar zenith angle is equal to 90° minus the solar elevation angle (From Lillesand et al. 2004).

Lillesand et al. (2004) mentioned that the earth-sun distance correction is applied to normalize for the seasonal changes in the distance between the Earth and the Sun. The earth-sun distance is usually expressed in the astronomical units. (An astronomical unit is equivalent to the mean distance between the Earth and the Sun, approximately 149.6 x 10^6 km.) The irradiance from the Sun decreases as the square of the earth-sun distance. If we ignore atmospheric effects, the combined influence of solar zenith angle and earth-sun distance on the irradiance incident on the Earth's surface can be expressed as:

$$E = \frac{E_0 \cos\theta_0}{d^2} \qquad (1\text{-}19)$$

where:

E is normalized solar irradiance,

E_0 is solar irradiance at mean earth-sun distance,

θ_0 is sun's angle from the zenith,

d is earth-sun distance, in astronomical units

1.8.1 The interactions of electromagnetic radiation with vegetation

The complex assemblage of the earth surface consists of a biological and physical feature. Both features involve directly in the remote sensing system. Vegetation is the one of biological features of the earth's feature that plays an important role in remote sensing system. Therefore, a fundamental understanding of electromagnetic radiation with vegetation, especially the spectral reflectance of vegetation and its effecting factor is important.

Knipling (1970) concluded that the reflectance of a plant canopy is similar, but is modified by the nonuniformity of incident solar radiation, plant structure, leaf areas, shadows, and background reflectivity. Sensors receive an integrated view of all these affects, and each crop or vegetation type tends to have a characteristics signature which permits its discrimination.

Colwell (1974) suggested the main factors that can be very important influencing vegetation canopy reflectance in certain situation: (a) leaf hemispherical reflectance and transmittance, (b) leaf area, (c) leaf orientation, (d) hemispherical reflectance and transmittance of supporting structures (stalks, trunks, limbs, petioles), (f) solar zenith angle, (g) look angle, and (h) azimuth angle.

Hoffer (1987) summarized distinct difference in reflectance of vegetation and their effecting factors that are found among the visible, near-infrared, and middle-infrared portions of the spectrum. In visible wavelengths, the pigmentation of the leaves is the dominanting factor. Most of incident energy is absorbed and remainder is reflected. The internal structure of the leaves controls the level of reflectance of the near infrared, where about half of the incident energy is reflected, nearly half is transmitted, and very little is absorbed by the leaf. The total moisture content of the vegetation controls the middle-infrared reflectance, with much of the incident energy being absorbed by the water in the leaf, remainder being reflected. Characteristic of spectral reflectance curve of vegetation is shown in Figure 1-24.

Furthermore, Curran (1985) confined the spectral reflectance of vegetation and its effect factor into two different types of reflectance.

1. The hemispherical reflectance of vegetation. By definition, the hemispherical reflectance infers the angles of incidence and collection of reflectance are hemispherical. It is usually operated in laboratory. Each of three features of leave: pigmentation, physiological structure and water content have an effect on the reflectance, absorbance and transmittance properties of a green leaf as indicated in Figure 1-24.

2. The bidirectional reflectance of vegetation. By definition, bidirectional reflectance infers the angles of incidence and collection of reflectance are directional as would be the case with satellite sensor measurements of radiance on a sunny day. With constant hemispherical reflectance of the individually leaves the bidirectional reflectance could vary appreciably due to effect of the soil background, the presence of senescent vegetation, the angular elevation of the sun and sensor, the canopy geometry and certain episodic and phenological canopy changes.

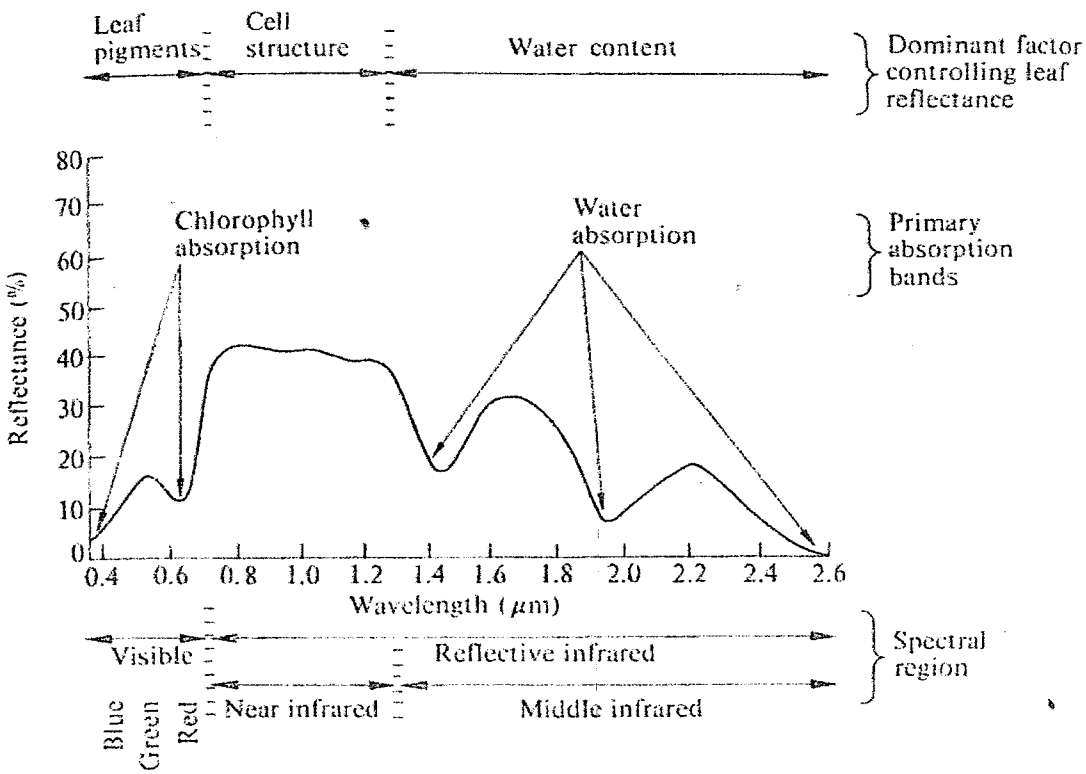


Figure 1-24: Significant spectral responses characteristics of green vegetation (From Hoffer, 1987).

Basically, Tucker (1987) identified spectral regions between 0.35 and 2.50 μm where different physiological variables control the resulting canopy spectral reflectance, which are important for the interpretation of remote sensing data on vegetation as following:

1. The 0.35 - 0.50 μm region is characterized by strong absorption by carotenoids and chlorophylls. At this spectral region relationship exists between spectral reflectance and the plant pigments present.

2. Between 0.50-0.62 μm a reduced level of pigment absorption is found. This results in a weaker relationship between spectral reflectance and the plant material.

3. The 0.62-0.70 μm region is characterized by a strong chlorophyll absorption, and high correlation exists between spectral reflectance and the chlorophyll concentration in the vegetation.

4. The 0.70-0.74 μm region is characterized by the transition from strong chlorophyll absorption and the high reflectance of green vegetation starting at around 0.75 μm. Consequently, there is a low relationship between the amount of green vegetation and reflectance.

5. Between 0.74-1.10 μm high levels of reflectance occur in the absence of major absorption of water in the vegetation. At this spectral interval a strong relation exists between the reflectance and the amount of water in the leaves.

Summary of the spectral region that are useful monitoring vegetation is given in Table 1-3.

Table 1-3: Ordered list of spectral regions in descending usefulness for monitoring green vegetation

Number	Wavelength(μm)	Utility for vegetation
1	0.74-~1.1	Direct biomass sensitivity
2	0.63-0.69	Direct in vivo chlorophyll sensitivity
3	~1.30-2.50	Direct in vivo foliar water sensitivity
4	0.37-0.50	Direct in vivo carotenoid and chlorophyll sensitivity
5	0.50-0.62	Direct/indirect and slight sensitivity to chlorophyll
6	0.70-0.74	Indirect and minimal sensitivity to vegetation
7	~1.10-1.30	Perhaps valuable non-vegetational information

From Tucker, 1987

1.8.2 The interactions of electromagnetic radiation with water

Energy incident on a water surface is primarily reflected from or transmitted into the surface, where the properties of each depend upon the angle of incident between the incident energy and the surface. The surface reflected energy is specular in character, usually as grazing rays from wave elements. Most of the energy is transmitted into the water body where it is absorbed by the chemical constituents of the water body, scattered by particulate matter in the water or reflected from the floor of the water body (McCloy, 2006).

Jensen (2007) described that the total radiance, (L_t) recorded by a remote sensing system over a water body is a function of the electromagnetic energy from four sources (Figure 1-25) as:

$$L_t = L_p + L_s + L_v + L_b \quad (1\text{-}20)$$

where:

- L_p is the radiance recorded by a sensor resulting from the downwelling solar (Esun) and sky (Esky) radiation. This is unwanted *path radiance* that never reaches the water.
- L_s is the radiance that reaches the air-water interface *(free-surface layer* or *boundary layer*) but only penetrates it a millimeter or so and is then reflected from the water surface. This reflected energy contains spectral information about the near-surface characteristics of the water.
- L_v is the radiance that penetrates the air-water interface, interacts with the organic/inorganic constituents in the water, and then exits the water column without encountering the bottom. It is called *subsurface volumetric radiance* and provides information about the internal bulk characteristics of the water column.

- L_b is the radiance that reaches the *bottom* of the waterbody, is reflected from it and propagates back through the water column, and then exits the water column. This radiance is of value if we want information about the bottom (e.g., depth, color).

The absorptance of water increases with increasing wavelength from a minimum absorptance at about 420 nm until it is greater than 90% in the reflected infrared portion of spectrum (Figure 1-26). Most scattering and reflection from water is therefore in the blue-green portion of spectrum, particularly for relatively pure water. Both chlorophyll and suspended soil sediment significantly affect the reflectance of water, with both of them absorbing energy in accordance with their individual absorptance characteristics (McCloy, 2006). The effect of chlorophyll on absorptance and the effects of some types of sediment are shown in Figure 1-27. The effects of sediments depend on the reflectance characteristics of the sediment particle, the density of the particle in the water and the depth of the particulate layer.

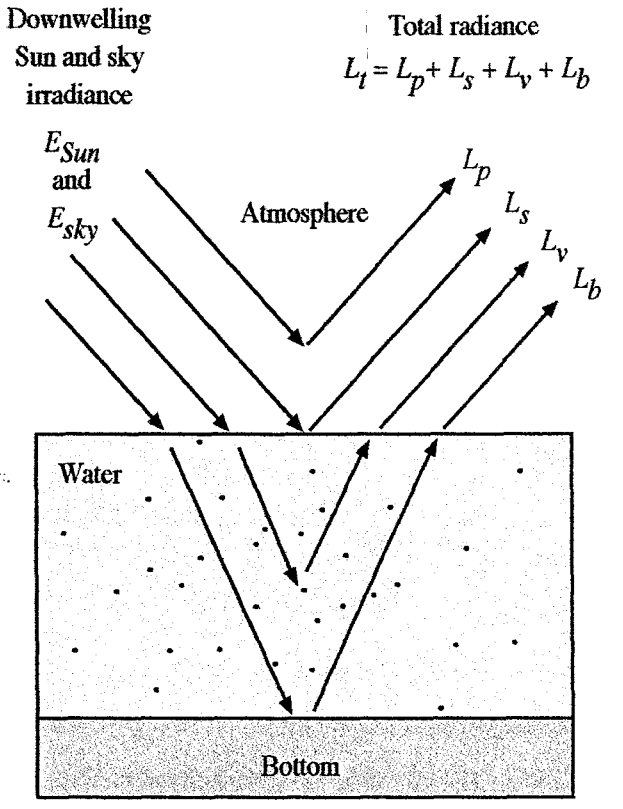


Figure 1-25: Water bodies receive irradiance from the Sun and atmosphere (From Jensen, 2007).

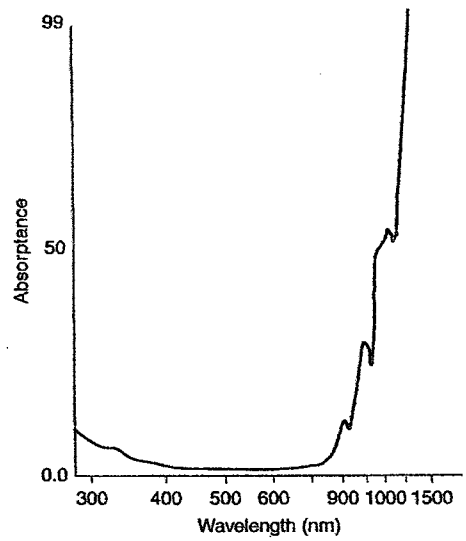


Figure 1-26: The absorption characteristics of pure water (From McCloy, 2006).

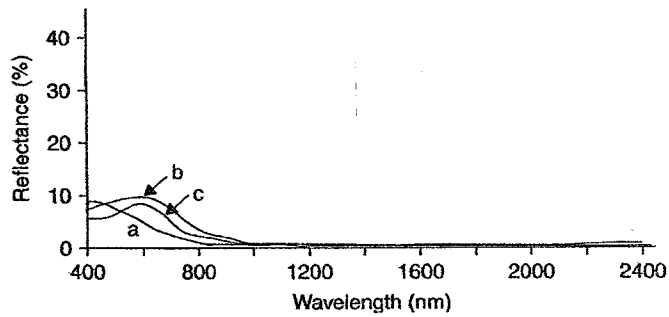


Figure 1-27: The reflectance characteristic of variant water: (a) pure water, (b) the effects of sediments water and (c) chlorophyll in the water (From McCloy, 2006).

Jensen (2007) mentioned that when conducting water-quality studies using remotely sensed data, we are usually most interested in measuring the *subsurface volumetric radiance*, L_v exiting the water column toward the sensor. The characteristics of this radiant energy are a function of the concentration of pure water (*w*), inorganic suspended minerals (*SM*), organic chlorophyll *a* (*Chl*), dissolved organic material (*DOM*), and the total amount of absorption and scattering attenuation that takes place in the water column due to each of these constituents, $c(\lambda)$:

$$Lv = f[wc(\lambda), SMc(\lambda), Chlc(\lambda), DOMc(\lambda)] \quad (1\text{-}21)$$

It is useful to look at the effect that each of these constituents has on the spectral reflectance characteristics of a water column.

1.8.3 The interactions of electromagnetic radiation with soil

Theoretically, the total upwelling radiance from an exposed soil recorded by the sensor, L_t, onboard the aircraft or satellite, L_p, is a function of the electromagnetic energy from the sources identified (Jensen, 2007) as shown in Figure 1-28 and summarized as:

$$L_t = L_p + L_s + L_v \quad (1\text{-}22)$$

where:

- L_p is the radiance recorded by a sensor resulting from the downwelling solar (Esun) and sky (Esky) radiation. This is unwanted *path radiance* that never reaches the soil surface. This is unwanted atmospheric path radiance noise and should ideally be removed from the imagery prior to trying to extract information about surficial soils or minerals.
- L_s is the radiance that reaches the air-soil interface *(boundary layer)* and penetrates it approximately 1/2 wavelength deep into the soils. The amount of radiant flux exiting the soil column based on the reflection and scattering taking place at this depth is L_s. The characteristics of the soil organic matter (decomposed vegetation or animal material) and inorganic (mineral) constituents and the amount of soil moisture present have a significant impact on the amount of energy absorbed, scattered, and/or reflected from the surficial portion of the soil/rock matrix (Figure 28a, 28b and 28c).
- L_v is volume scattering. Some of the incident downwelling solar and sky radiation may be able to penetrate perhaps millimeters or even a centimeter or two into the soils column. In fact, almost all the specular boundary layer reflectance and volumetric scattering may take place in the first few millimeters of the soil profile. The amount of volumetric radiant flux scattered or reflected back into the atmosphere is a function of the wavelength of incident energy, the type and amount of organic/inorganic constituents, the shape and density of the minerals, the degree of mineral compaction, and the amount of soil moisture present.

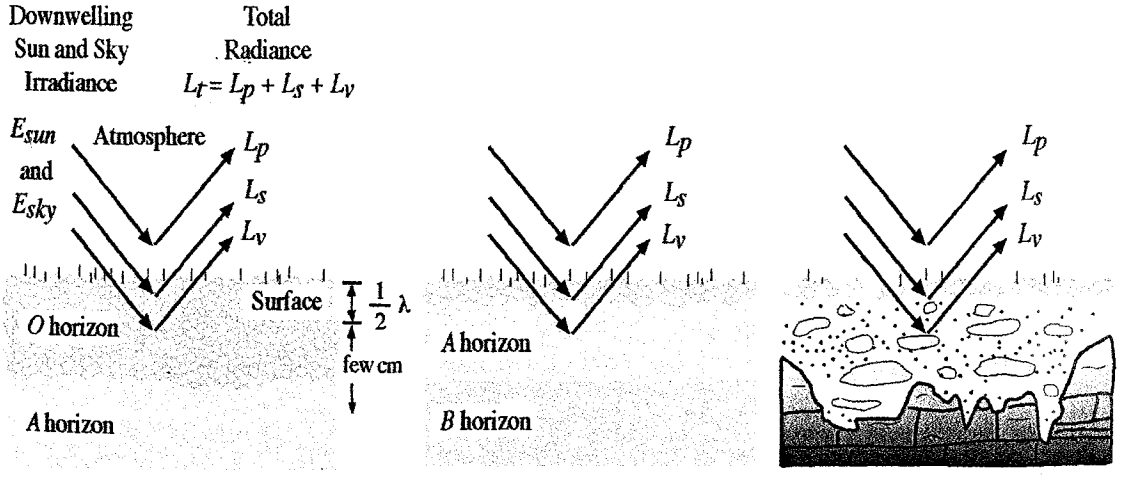


Figure 1-28: Soils and rocks receive irradiance from the Sun and atmosphere (From Jensen, 2007)

Soil reflectance occurs in the surface layers of the soil, with a significant proportion occurring as surface reflectance from the soil grains themselves. McCloy (2006) claimed that affect the reflectance of soils were:

1. Chemical composition. The chemical composition of the soil affects soil color due to selective absorption. Difference chemicals have different absorption bands and in consequence the existence of different chemicals in the soil will affect the selective absorption of energy during transmission through the soil. The most obvious effect of this type is the high red reflectance of soil rich in ferric iron.

2. Humus content. Humus, the organic products derived from the breakdown or decomposition of vegetables and animal matters, usually exhibits absorption at all wavelengths in the visible region, but with slightly higher absorption in the blue-green part of the spectrum due to the antho-cyanin compounds created by the breakdown of chlorophyll in leaf elements. The level of humus in soils varies over time so that the reflectance of soils can also vary over time. The decompositional state of plant litter significantly affects the reflectance curves of organic soils. Characteristics of soil spectral reflectance with three variants decompositional state is shown in Figure 1-29.

3. Surface soil moisture. Water selectively absorbs at all wavelengths, but with increasing absorptance at longer wavelengths. Since, most of the reflectance of soils occurs in the top layers of soil particles, only the water in these top layers will affect the reflectance of the soil surface. If these layers are moist then the water will have a strong influence of surface water on soil reflectance is quite minimal. The influence of these levels of moisture content on the reflectance of soil type is shown in Figure 1-30.

4. Surface roughness. The combination of surface roughness and solar elevation affects the amount of shadowing that occurs in the soil and in consequence the soil reflectance. Surface roughness can be due to the soil grain size for soils and sands that do not form clods, or the soil texture due to the clod size for other soils. For example, claypans or flat surfaces with a very finely textured surface are more highly reflective than rougher sandy surfaces even though sand itself is often more reflective than the particles of clay in the claypan. In addition, surface roughness can also be due to cultivation or other human activities on the soil surface that break up the surface and either increase or decrease its texture. Cultivation increases the surface clod size immediately after cultivation, thereby increasing the texture and reducing the reflectance due to increased shadowing by the clods.

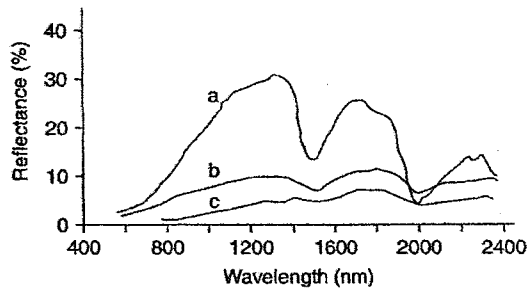


Figure 1-29: Representative spectra for the reflectance of soils with (a) minimally altered (fibric), (b) partially altered (ferric) and (c) fully decomposed (sapric) organic matter (From McCloy, 2006).

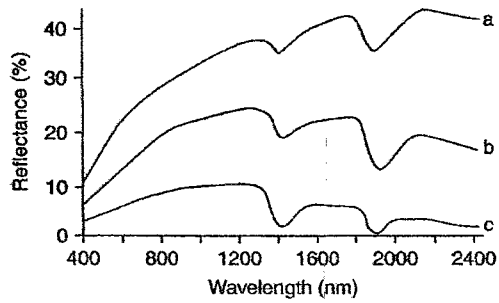


Figure 1-30: Reflectance of a typical soil with changes in moisture content: (a) dry (b) normal and (c) moist conditions (From McCloy, 2006)

Furthermore, Jensen (2007) mentioned that the spectral reflectance characteristic of soils are a function of several important characteristics, included soil texture (percentage of sand, silt and clay), soil moisture content (e.g., dry, moist, saturated), organic mater content, iron-oxide content, soil salinity, and surface roughness.

Finally, Heute (2004) also mentioned that the spectral composition of soil-reflected and emitted energy is primarily dependent on the biogeochemical (mineral and organic) constituents, optical-geometric scattering (particle size, aspect, roughness), and moisture conditions of the surface. The aim of remote sensing is to exploit and model these complex relationships and pattern of surface energy interactions for the purpose of mapping and extracting information about the biophysical and biochemical character of soils.

1.9 Radiometric Concepts, Terminology and Units.

1.9.1 Radiometric Characteristics of Radiation Measurement

The characteristics of radiant flux and what happens to it as it interacts with the Earth's surface is of critical importance in remote sensing. In fact, this is the fundamental focus of much remote sensing research (Jensen, 2005). Lillesand et al. (2004) stated that in most applications of remote sensing, radiation is measured in radiometric units. Which particular terms and units are used in any given situation is often determined by the angular nature of the measurement

involved. Figure 1-31 illustrated the two broad categories of angular measurement commonly used in remote sensing. There are (a) hemispherical measurement and (b) directional measurement.

- Hemispherical measurements account for all energy contained within a hemisphere above a surface or object of interest. The term *irradiance* refers to the hemispherical radiation that is incident on an object or area. The hemispherical radiation that leaves an object or area is described by the term *radiant exitance*.
- Directional radiation measurements account for the radiation measured in a particular direction of illuminating or view. An importance aspect of directional measurement is the cosine effect. This concept is illustrated in Figure 1-32 for an object that does not fill the field of view of an optical system. When such an object is viewed in a perpendicular, or nominal direction, it will appear to optical system to have a certain area (A).

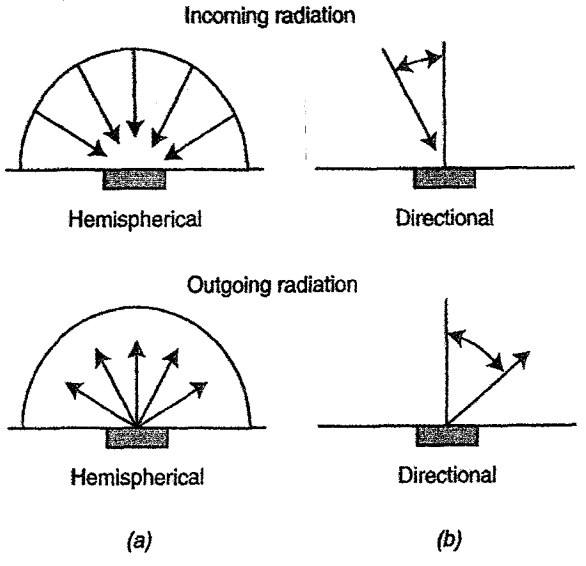


Figure 1-31: Measurement of incoming and outgoing radiation (a) Hemispherical and (b) directional (From Lillesand et al, 2004)

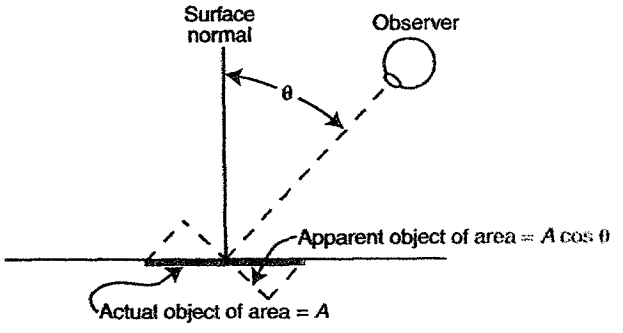


Figure 1-32: Projected area cosine in a viewing direction other than normal (From Lillesand et al., 2004)

Also, central to understand directional radiation measurement are concepts of plane-angle and solid-angle measurements (Lillesand et al. (2004). These two forms of angular measurement are displayed in Figure 1-33.

- Plane angle radiation measurement (Figure 1-33a) is a plane angle that is measured in two dimensions and usually expressed in *radians*. The size of a plane angle (β) expressed in radians is equal to the length of the arc of a circle subtended by the angle divided by the radius of the circle (when the arc length exactly equals the radius, the size of the angle is exactly 1 radian).
- Solid angle radiation measurement (Figure 1-33b) is a solid angle that is measured in three dimensions and express in units of steradians. This size of a solid angle (Ω) expressed in steradians is equal to the area subtended by the surface of a sphere divided by the square of the radians of the sphere (when the subtended spherical area exactly equals the square of the radius, the size of the solid angle is 1 steradian).

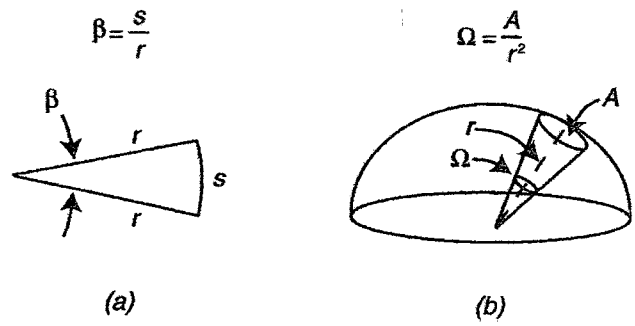


Figure 1-33: Defining angle used in radiation measurement (a) plane angle and (b) solid angle (From Lillesand et al., 2004).

1.9.2 Radiometric Terminology and Units

Lillesand et al (2004) illustrated the basic relationship among the various terms that were used in both hemispherical and directional radiation measurement as shown in Figure 1-34. The definition of selected radiation based on Lillesand et al (2004) and quantitative measurement based on Jensen (2005) can be summarized in Table 1-4 and briefly described as follows:

Radiant energy (Q)

Radiant energy is the energy carried by an electromagnetic wave and measure of the capacity of the wave to do work, measured in joules.

Radiant flux (Φ)

Radiant flux is the amount of radiant energy emitted, transmitted, or received per unit time, measured in watts (joules per second).

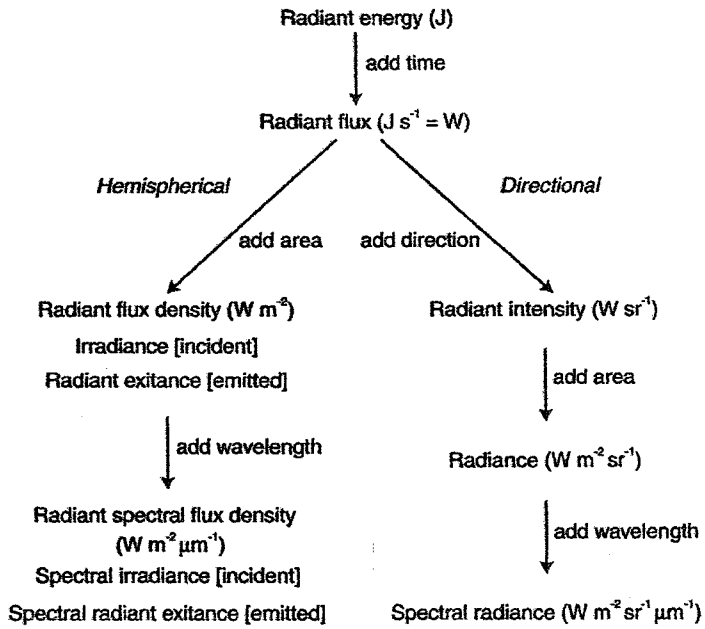


Figure 1-34: Relationship among the various terms in hemispherical and directional radiation measurement (From Lillesand et al., 2004).

Radiant flux density

Radiant flux density is the radiant flux at a surface divided by the area of the surface. Again, the density for flux incident upon a surface is *irradiance*, (*E*). The density for flux leaving a surface is *radiant exitance* (*M*). Both are measured in watt per square meter.

Radiant spectral flux density

Radiant spectral flux density is the radiant flux density per unit wavelength interval. Both *spectral irradiance*, (E_λ) and *spectral radiant exitance* (M_λ) as shown in Figure 1-35 are measured in units of watts per square meter per micrometer.

Also, both are calculated as follows:

Spectral irradiance: $$E_\lambda = \frac{\Phi_\lambda}{A} \quad (1\text{-}23)$$

Spectral radiant exitance: $$M_\lambda = \frac{\Phi_\lambda}{A} \quad (1\text{-}24)$$

Radiant Intensity (I)

Radiant intensity is the flux emanating from a point source per unit solid angle in the direction considered, measured in watt per steradian.

Radiance (L)

Radiance is the radiant flux per unit solid angle emanating from a surface in a given direction per unit of projected surface in the direction considered. Figure 1-36 illustrated the concept of radiance. Radiance is measured in units of watts per square meter per steradian.

Spectral radiance (L_λ)

Spectral radiance is the radiance per wavelength interval, measure in watts per square meter per steradian per micrometer. The spectral radiance was computed by:

Spectral radiance: $$L_\lambda = \frac{\frac{\Phi_\lambda}{\Omega}}{A\cos\theta} \quad (1\text{-}25)$$

Table 1-4: Radiometric Term, Symbol, Measure, Units

Term	Symbol	Defining Expression	Units
Radiant energy	Q		Joules, J
Radiant flux	Φ	$\frac{dQ}{dT}$	Watts, W
Radiant flux density: Irradiance	E	$\frac{d\Phi}{dA}(in)$	Watt per square meter, W m^{-2}
Radiant flux density: Radiant exitance	M	$\frac{d\Phi}{dA}(out)$	Watt per square meter, W m^{-2}
Radian spectral flux density: Irradiance	E_λ	$\frac{dE}{dL}(in)$	Watt per square meter per micrometer, W $m^{-2}\mu m^{-1}$
Radian spectral flux density: Radiant exitance	M_λ	$\frac{dE}{dL}(out)$	Watt per square meter per micrometer, W m^{-2} μm^{-1}
Radiant intensity	I		Watt per Steradian, W sr^{-1}
Radiance	L	$\frac{d^2\Phi}{\cos\theta dA d\omega}(out)$	Watt per square meter per steradian, W m^{-2} sr^{-1}
Spectral radiance	L_λ	$\frac{dL}{d\lambda}$	Watt per square meter per steradian, W m^{-2} sr^{-1} μm^{-1}

From Lillesand et al (2004) and Landgrebe (2003)

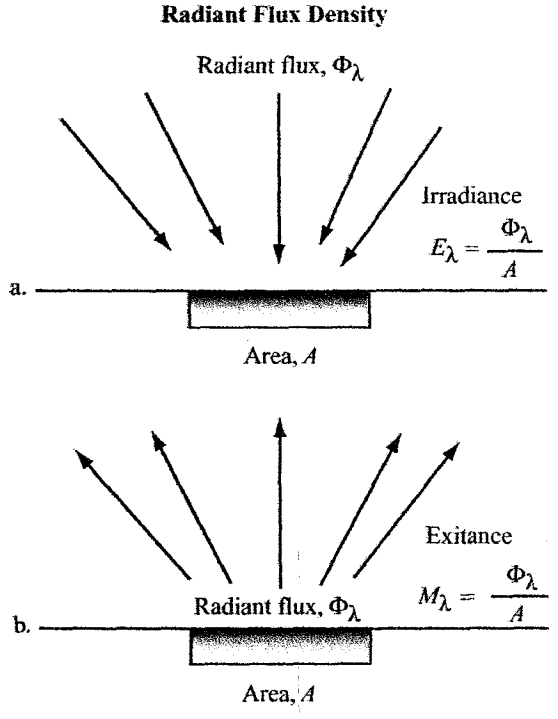


Figure 1-35: Characteristics of radiant flux density (From Jensen, 2005)

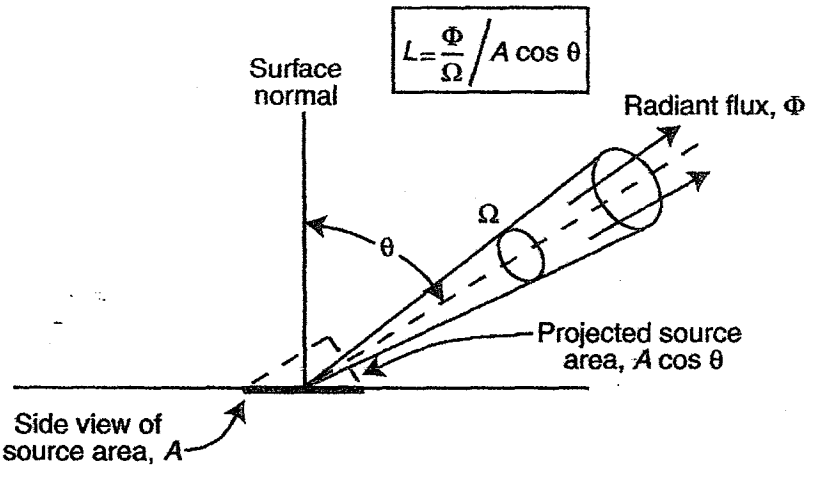


Figure 1-36: Concept of radiance (From Lillesand et al., 2004).

Chapter 2: Sensor in Remote Sensing

Sabins (1987) classified sensor system in remote sensing into two categories: framing system and scanning system. Characteristics of each system are described as following:

2.1 Framing system.

Framing system instantaneously acquires an image of an area, or frame, on the terrain. Camera and vidicons are common examples of such system (Figure 2-1). A camera employs a lens to form an image of the scene at the focal plane, which is the plane at which the image is sharply defined. A shutter opens at selected intervals to allow light to enter the camera, when the image is recorded on photographic film. A vidicon is a type of television camera that records the image on a photosensitive electronically charged surface. An electron beam then sweeps the surface to detect the pattern of charge differences a signal that may be transmitted and recorded on magnetic tape for eventual display on film.

Successive frames of camera and vidicon images may be acquired with forward overlap, as shown in Figure 2-1. The overlapping portion may be viewed with a stereoscope to produce a three-dimensional view. Film is sensitive only to portions of the ultraviolet, visible, and reflected infrared regions (0.3 to 0.9 μm). The sensitivity range of special vidicons extends into the thermal band of the infrared region. A framing system can instantaneously image a large area because the system has a dense array of detectors located at the focal plane. The emulsion of camera film contains tiny grains of silver halide. A vidicon surface is coated with sensitive phosphors.

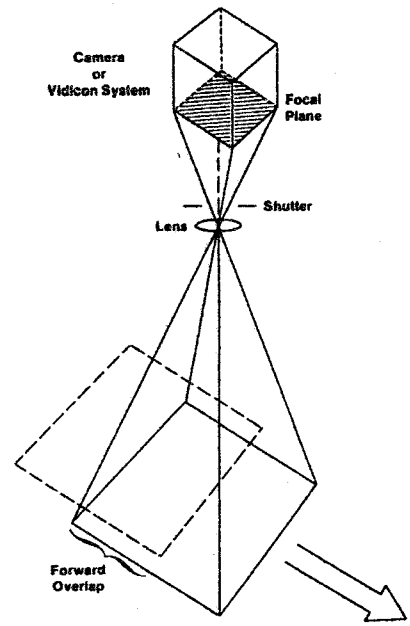


Figure 2-1: Framing system to acquiring remote sensing images (From Sabins, 1987)

2. 2 Scanning system

A scanning system employs a single detector with a narrow field of view which sweeps across the terrain to produce an image. When photons of electromagnetic energy radiated or reflected from the terrain encounter the detector, an electrical signal is produced that varies in proportion to the number of photons. The electrical signal is amplified, recorded on magnetic tape, and played back later to produce an image. All scanning systems sweep the detector's field of view the terrain in a series of parallel scan lines. There are four common scanning models: *cross-track scanning, circular scanning, along-track scanning, and side scanning.*

2.2.1 Cross-Track Scanning System

The widely used cross-track scanning systems employ a faceted mirror that is rotated by an electric motor, with a horizontal axis of rotation aligned parallel with the flight direction (Figure 2-2). The mirror sweeps across the terrain in a pattern of parallel scan lines oriented normal (perpendicularly) from the ground is focused onto the detector by secondary mirrors.

The angular resolving power of a detector, measured in milliradians, determines the Instantaneous Field of View (IFOV) of the detector. As shown in Figure 2-2, the IFOV subtends an area on the terrain called a ground resolution cell. Dimensions of ground resolution are determined by the detector's IFOV and the altitude of the scanning system. A detector with an IFOV of 1 mryd at an altitude of 10 km has a ground resolution all of 10 by 10 m.

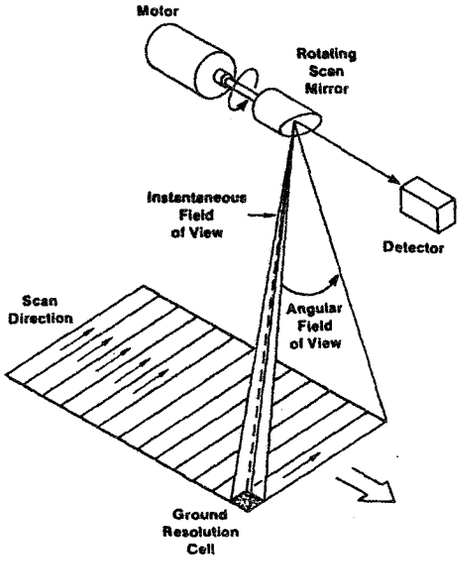


Figure 2-2: Cross-Track Scanner (From Sabins, 1987)

The angular field of view (Figure 2-2) is that portion of the mirror sweep, measured in a degree, is recorded as a scan line. The angular field of view and the altitude of the system determine the ground swath, which is the width of the terrain strip represented by the image.

2.2.2 Circular Scanning System

In a circular scanning system, the scan motor and mirror are mounted with a vertical axis of rotation that sweeps a circular path on the terrain (Figure 2-3). Only the forward portion of the sweep is recorded to produce images. An advantage of this system is that the distance between scanner and terrain is constant and all the ground resolution cells have the same dimensions. The major disadvantage is that most image processing and display systems are designed for linear scan data, therefore the circular scan data must be extensively reformatted prior to processing. Circular scanners have short dwell times comparable to those of cross-track scanners. Circular scanners are used for reconnaissance purposes in helicopters and low-flying aircraft. The axis of rotation is tilted to point forward and acquired images of the terrain well in advance of aircraft position. The images are displayed in real time on a screen in the cockpit to guide the pilot.

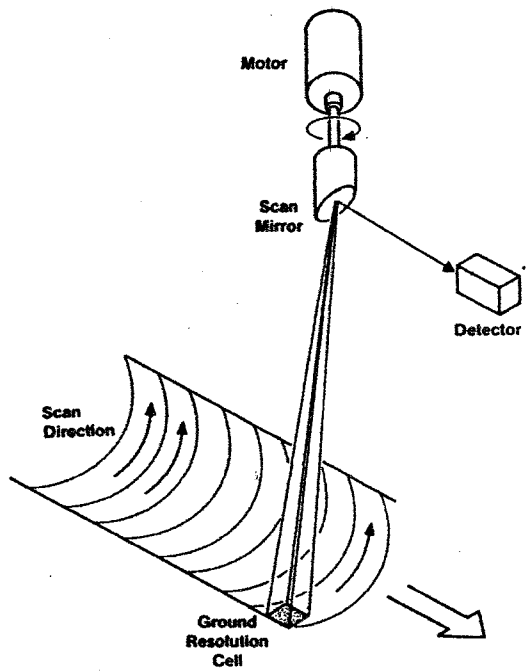


Figure 2-3: Circular Scanner (From Sabins, 1987)

2.2.3 Along-Track Scanning System.

For scanner systems to achieve finer spatial and spectral resolution, the dwell time for each ground resolution cell must be increased. One method is to eliminate the scanning mirror and provide an individual detector for each ground resolution cell across the ground swath (Figure 2-4). The detectors are placed in a linear array in the focal plane of the image formed by a lens system.

The long axis of the linear array is oriented normal to the flight path, and the IFOV of each detector sweeps a ground resolution all along the terrain parallel with the flight track direction. Along-track scanning refers to this movement of the ground resolution cells. These systems are also called pushbroom scanner because the detectors are analogous to the bristles of a broom pushed along the floor.

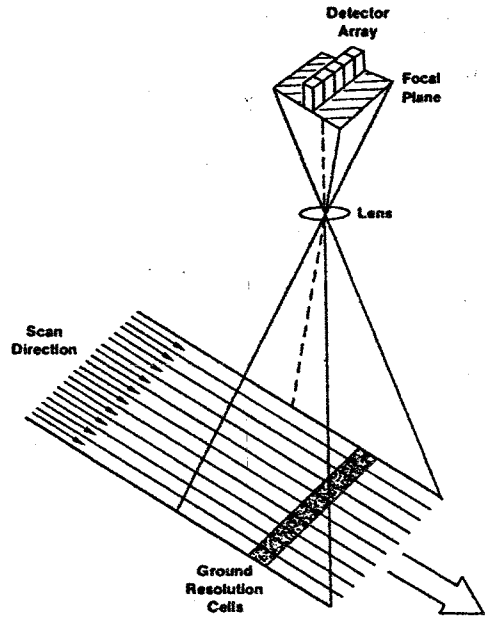


Figure 2-4: Along-Track Scanner (From Sabins, 1987)

2.2.4 Side Scanning System

The three types of scanners described above are passive systems, since they detect and record energy naturally reflected or radiated from the terrain. Active systems, which provide their own energy source, operate primarily in the side-scanning model. The example in Figure 2-5 is a radar system that transmits pulse of microwave energy to one side of the flight path (range direction) and records the energy scattered from the terrain back to the antenna. Another system is side-scanning sonar, which transmits pulse of sonic energy in the ocean to map bathymetric features.

The strength of the signal generated by a detector is a function of the following factors:

Energy flux. The amount of energy reflected a radiant from terrain is the energy flux. For visible detectors, this flux is lower on a dark day than on a sunny day.

Altitude. For a given ground resolution cell, the amount of energy reaching the detector is inverse proportional to the square of the distance. At greater altitudes the signal strength is weaker.

Spectral bandwidth of the detector. The signal is stronger for detectors that respond to a broader-wavelength range of energy. For example, detector that is sensitive to the entire visible range will receiver more energy than a detector that is sensitive to a narrow band, such as visible red.

Instantaneous field of view. Both the physical size of the sensitive element of the detector and the effective focal length of the scanner optics determine the IFOV. A small IFOV is required for high spatial resolution but restricts the signal strength (amount of energy receive by the detector).

Dwell time. The time required for the detector IFOV to sweep across a ground resolution all is the dwell time. Along dwell time allows more energy to impinge on the detector, which creates a stronger signal.

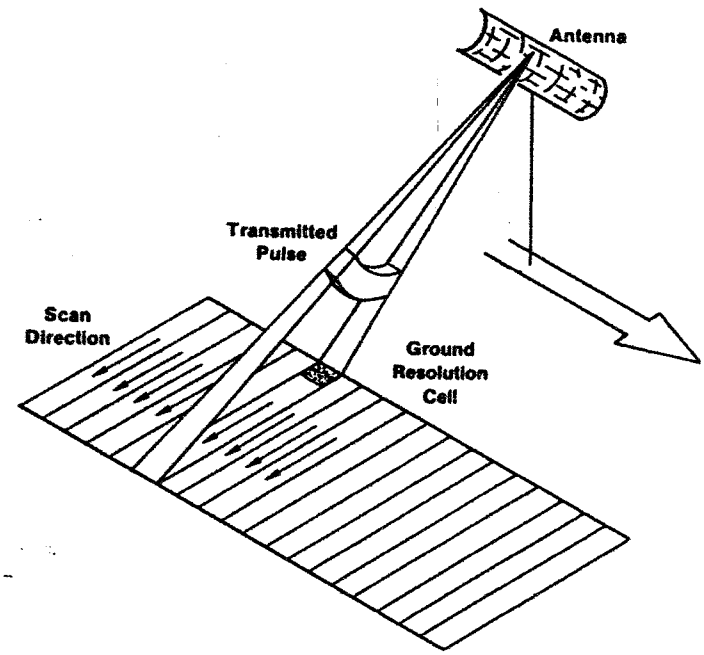


Figure 2-5: Side Scanning System (From Sabins, 1987)

Jensen (2007) categorized the digital multispectral and hyperspectral remote sensing systems based the type remote sensing technology used as summarized in Figure 2-6, including:

- traditional aerial photography;
- multispectral imaging using a scanning mirror and discrete detectors;
- multispectral imaging with linear arrays;
- imaging with a scanning mirror and linear arrays;
- imaging spectrometry using linear and area arrays;
- digital frame cameras aerial photography based on area arrays.

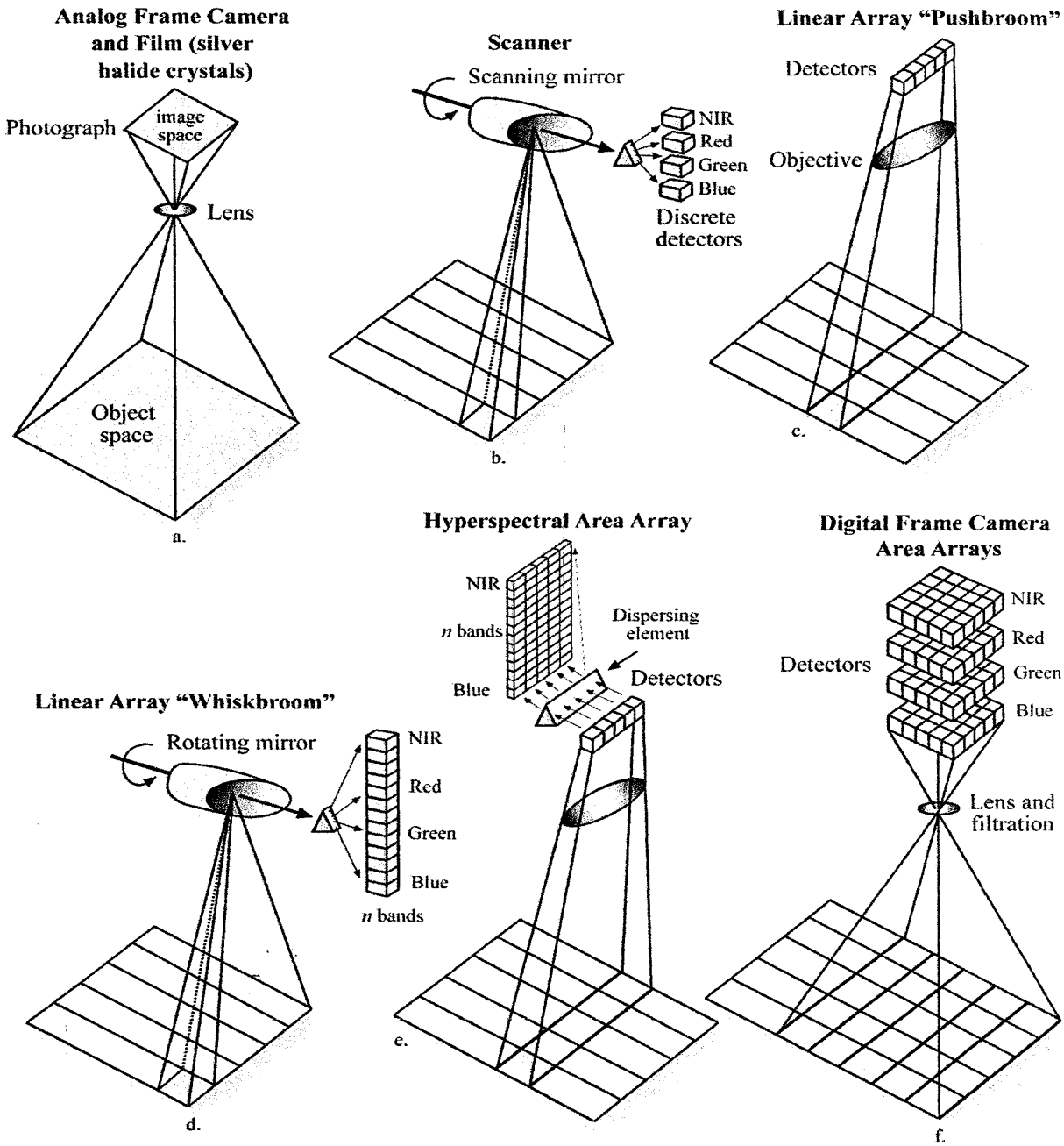


Figure 2-6: Six types of remote sensor systems used for multispectral and hyperspectral data collection: a) traditional aerial photography, b) multispectral imaging using a scanning mirror and discrete detectors, c) multispectral imaging with linear arrays d) imaging with a scanning mirror and linear arrays, e) imaging spectrometry using linear and area arrays, and f) digital frame cameras aerial photography based on area arrays. (From Jensen, 2007).

A. Multispectral Imaging Using Discrete Detectors and Scanning Mirrors

- Landsat Multispectral Scanner (MSS)
- Landsat Thematic Mapper (TM)
- Landsat Enhanced Thematic Mapper Plus (ETM+)
- NOAA Geostationary Operational Environmental Satellite (GOES)
- NOAA Advanced Very High Resolution Radiometer (AVHRR)
- NASA and ORBIMAGE, Inc., Sea-Viewing Wide field-of –view Sensor (SeaWiFS)
- Daedalus, Inc., AirCraft Multispectral Scanner (AMS)
- NASA Airborne Terrestrial Applications Sensor (ATLAS)

The major characteristics of selected multispectral imaging using discrete detectors and scanning mirrors were summarized in Table 2-1 to Table 2-4, namely, Landsat, NOAA, SeaWiFS, AMS and ATLAS, respectively.

B. Multispectral Imaging Using Linear Arrays

- SPOT 1, 2 and 3 High Resolution Visible (HRV) sensors
- SPOT 4 and 5 High Resolution Visible Infrared (HRVIR) and Vegetation sensor
- Indian Remote Sensing System (IRS) Linear Imaging Self-scanning Sensor (LISS-III and LISS-IV)
- NASA *Terra* Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER)
- NASA *Terra* Advanced Spaceborne Thermal Emission and Reflection Radiometer (MISR)
- Digital Globe, Inc. (Quickbird)
- Space Imaging, Inc. (IKONOS)
- ImageSat International, Inc. (EROS A1)
- ORBIMAGE, Inc. (OrbView-3 and OrbView-5)
- Leica Geosystems, Inc. Airborne Digital Sensor System (ADS-40)

The major characteristics of selected multispectral imaging using linear arrays were summarized in Table 2-5 to Table 2-10, namely, SPOT, IRS, ASTER, IKONOS/OrbView-3/QuickBird, THEOS, and ADS-40, respectively.

C. Imaging Spectrometry Using Linear and Area Arrays

- NASA Jet Propulsion Laboratory Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)
- Compact Airborne Spectrographic Imager 1500 (CASI 1500) and Hyperspectral SWIR sensor (SASI 600)
- NASA Terra Moderate Resolution Imaging Spectrometer (MODIS)
- NASA Earth Observation (EO-1) Advance Land Imager (ALI), Hyperion, and LEISA Atmospheric Corrector (LAC)

The major characteristics of selected imaging spectrometry using linear and area arrays were summarized in Table 2-11 to Table 2-12, namely, AVIRIS and CASI 3, and MODIS, respectively.

D. Digital Frame Cameras Based on Area Arrays

- Leica Geosystems Emerge Digital Sensor Data (DSS)
- Vexcel UltraCam Large Format Camera
- Z/1 Digital Modular Camera

The major characteristic of selected digital frame cameras based on area arrays of DSS was summarized in Table 2-13.

E. Astronaut Photographic Systems

- NASA Space Shuttle and International Space Station Imagery.

Table 2-1: ***Landsat Multispectral Scanner (MSS) Landsat Thematic Mapper (TM) and Landsat Enhanced Thematic Mapper Plus (ETM+)***

LANDSAT 1-2-3			LANDSAT-4-5			LANDSAT 6-7		
Band	Spectral Resolution (μm)	Spatial Resolution (m)	Band	Spectral Resolution (μm)	Spatial Resolution (m)	Band	Spectral Resolution (μm)	Spatial Resolution (m)
4	0.5 – 0.6	79 x 79	1	0.45 – 0.52	30 x 30	1	0.450– 0.515	30 X 30
5	0.6 – 0.7	79 x 79	2	0.52 – 0.60	30 x 30	2	0.525–0.605	30 X 30
6	0.7 – 0.8	79 x 79	3	0.63 – 0.69	30 x 30	3	0.630–0.690	30 X 30
7	0.8 – 1.1	79 x 79	4	0.76 – 0.90	30 x 30	4	0.750–0.900	30 X 30
8	10.4 – 12.6	240x240	5	1.55-1.75	30 x 30	5	1.55–1.75	30 X 30
			6	10.40-12.5	120x120	6	10.40–12.50	60 X 60
			7	2.08-2.35	30 x 30	7	2.08 – 2.35	30 X 30
						8 Pan.	0.52 – 0.90	15 X 15
Quantization levels		6 bit	Quantization levels		8 bit	Quantization levels		8 bit
Swath Width		185 km	Swath Width		185 km	Swath Width		185 km
Revisit		18 days	Revisit		16 days	Revisit		16 days
Altitude		919 km	Altitude		705 km	Altitude		705 km
Inclination		99 °	Inclination		98.2 °	Inclination		98.2 °

From Jensen (2007)

Table 2-2: NOAA Advanced Very High Resolution Radiometer (AVHRR) sensor system characteristics

Band	NOAA-6,8,10 Spectral Res. (μm)	NOAA-9,11,12,13,14 Spectral Res. (μm)	NOAA-15,16,17 AVHRR/3 Spectral Res. (μm) Res. (km)	Band Utility
1	0.580 – 0.68	0.580 – 0.68	0.580 – 0.68	Daytime clouds, snow, ice, and vegetation mapping: used to compute NDVI
2	0.725 – 1.10	0.725 – 1.10	0.725 – 1.10	Land/water interface, snow, ice and vegetation mapping: used to compute NDVI
3	3.55 – 3.93	3.55 – 3.93	3A: 1.58 – 1.64 3B: 3.55 – 3.93	Monitor hot targets (volcanoes, forest fires) night-time cloud mapping
4	10.50 – 11.50	10.30 – 11.30	10.30 – 11.30	Day/night cloud and surface-temperature mapping
5	None	11.50 – 12.50	11.50 – 12.50	Cloud and surface temperature, day and night cloud mapping; removal of atmospheric water vapor path radiance

From Jensen (2007)

Table 2-3: Characteristics of the Sea-viewing Wide Field-of-View Sensor (SeaWIFS)

Band	Band center (nm)	Bandwidth (nm)	Band Utility
1	412	402 – 422	Identify yellow substance
2	443	433 – 453	Chlorophyll concentration
3	490	480 – 500	Increased sensitivity to chlorophyll concentration
4	510	500 – 520	Chlorophyll concentration
5	555	545 – 565	Gelbstoffe (Yellow substance)
6	670	660 – 680	Chlorophyll concentration
7	765	745 – 785	Surface vegetation, land-water interface, atmospheric correction
8	865	845 – 885	Surface vegetation, land-water interface, atmospheric correction
Spatial Resolution: 1.13 X 1.13 km over a swath of 2800 km. Revisit time 1 day			

From Jensen (2007)

Table 2-4: System characteristics of the Daedalus Airborne Multispectral Scanner (AMS) and NASA Airborne Terrestrial Applications Sensor

Airborne Multispectral Scanner (AMS)			**NASA Airborne Terrestrial Application Sensor (ATLAS)**		
Band	**Spectral Resolution (μm)**	**Spatial Resolution (m)**	**Band**	**Spectral Resolution (μm)**	**Spatial Resolution (m)**
1	0.42 – 0.45	Variable depending upon altitude above ground level	1	0.45 – 0.52	2.5 to 25 m depending upon altitude above ground level
2	0.45 – 0.52		2	0.52 – 0.60	
3	0.52 – 0.60		3	0.60 – 0.63	
4	0.60 – 0.63		4	0.63 – 0.69	
5	0.63 – 0.69		5	0.69 – 0.76	
6	0.69 – 0.75		6	0.76 – 0.90	
7	0.76 – 0.90		7	1.55 – 1.75	
8	0.91 – 1.05		8	2.08 – 2.35	
9	3.00 – 5.50		9	Removed	
10	8.50 – 12.5		10	8.20 – 8.60	
			11	8.60 – 9.00	
			12	9.00 – 9.40	
			13	9.60 – 10.20	
			14	10.20 – 11.20	
			15	11.20 – 12.20	
IFOV	2.5 mrad		**IFOV**	2.0 mrad	
Quantization	8-12 bits		**Quantization**	8 bits	
Altitude	Variable		**Altitude**	variable	
Swath width	714 pixels		**Swath width**	800 pixels	

From Jensen (2007)

Table 2-5: SPOTs 1,2 and 3 High Resolution Visible (HRV), SPOTs 4 and 5 High Resolution Visible and Infrared (HRVIR), SPOTs 4 and 5 *Vegetation sensor system characteristics*

SPOT 1, 2, and 3			SPOT 4 and 4			SPOT Vegetation		
Band	**Spectral Res. (μm)**	**Spatial Res. (m) at nadir**	**Band**	**Spectral Res. (μm)**	**Spatial Res. (m) at nadir**	**Band**	**Spectral Res. (μm)**	**Spatial Res. (km) at nadir**
1	0.50 - 0.59	20 X 20	1	0.50 - 0.59	10 X 10	1	0.43 - 0.47	1.15 X 1.15
2	0.61 – 0.68	20 X 20	2	0.61 – 0.68	10 X 10	2	0.61 – 0.68	1.15 X 1.15
3	0.79 – 0.89	20 X 20	3	0.79 – 0.89	10 X 10	3	0.78 – 0.89	1.15 X 1.15
Pan	0.51 – 0.73	10 X 10	Pan	0.48 – 0.71	2.5 X 2.5			
Pan (4)	0.61 – 0.68	10 X 10						
SWIR (4)	1.58 – 1.75	20 X 20	SWIR	1.58 – 1.75	20 X 20	SWIR	1.58 – 1.75	1.15 X 1.15
Sensor	Linear array pushbroom		**Sensor**	Linear array pushbroom		**Sensor**	Linear array pushbroom	
Swath	60 km+/- 50.5 Deg.		**Swath**	60 km+/- 27 Deg.		**Swath**	2250 km+/- 50.5 Deg.	
Rate	25 Mb/s		**Rate**	50 Mb/s		**Rate**	50 Mb/s	
Revisit	26 days		**Revisit**	26 days		**Revisit**	1 days	
Orbit	822 km, Sun-synchronous, Inclination = 98.7 Deg., Equatorial crossing 10.30 a.m.		**Orbit**	822 km, Sun-synchronous, Inclination = 98.7 Deg., Equatorial crossing 10.30 a.m.		**Orbit**	822 km, Sun-synchronous, Inclination = 98.7 Deg., Equatorial crossing 10.30 a.m.	

From Jensen (2007)

Table 2-6: Indian National Remote Sensing Agency (NRSA) Indian Remote Sensing (IRS) satellite characteristics.

LISS-I and LISS-II Band	Spectral Res. (μm)	Spatial Res. (m) at nadir LISS-I	Spatial Res. (m) at nadir LISS-II	LISS-III PAN WiFS Band	Spectral Res. (μm)	Spatial Res. (m) at nadir
1	0.45 – 0.52	72.5	36.25	1	-	-
2	0.52 – 0.59	72.5	36.25	2	0.52 – 0.59	23.5 X 23.5
3	0.62 – 0.68	72.5	36.25	3	0.62 – 0.68	23.5 X 23.5
4	0.77 – 0.86	72.5	36.25	4	0.77 – 0.86	23.5 X 23.5
				5	1.55 – 1.70	70.5 X 70.5
				Pan	0.50 – 0.75	5.2 X 5.2
				WiFS-1	0.62 – 0.68	188 X 188
				WiFS-2	0.77 – 0.86	188 X 188
Sensor	Linear Array Pushbroom			Sensor	Linear Array Pushbroom	
Swath Width	LISS-I = 148 km; LISS-II = 146 km			Swath Width	LISS-III (B2/3/4) 141 km; B5 148 km; Pan 70km;WiFS 692 km	
Revisit	22 days at equator			Revisit	LISS-III 24 hr; Pan 5 days; WiFS 5 days	
Orbit	904 km, Sun-synchronous, Inclination = 99.5 Deg., Equatorial crossing 10.26 a.m.			Orbit	817 km, Sun-synchronous, Inclination = 98.69 Deg., Equatorial crossing 10.30 a.m. +/- 5 min	
Launch	IRS-1A: March 17,1988; IRS-1B: August 29,1991			Launch	IRS-1C: 1995; IRS-1B: September,1997	

From Jensen (2007)

Table 2-7: *NASA Advanced Spaceborne* Thermal Emission and Reflection Radiometer (ASTER) characteristics

Band	VNIR spectral Resolution (μm)	Band	SWIR spectral Resolution (μm)	Band	TIR spectral Resolution (μm)
1	0.52 – 0.60	4	1.600 – 1.700	10	8.125 – 8.475
2	0.63 – 0.69	5	2.145 – 2.185	11	8.475 – 8.825
3	0.76 – 0.86	6	2.185 – 2.225	12	8.925 – 9.275
3 backward	0.76 – 0.86	7	2.235 – 2.285	13	10.25 – 10.95
		8	2.295 – 2.365	14	10.95 – 11.65
		9	2.360 – 2.430		
Technology	Pushbroom	**Technology**	Pushbroom	**Technology**	Whiskbroom
Detector	Si	**Detector**	PtSi;Si	**Detector**	Hg:Cd:Te
Spatial Resolution	15 X 15 m	**Spatial Resolution**	30 X 30 m	**Spatial Resolution**	90 X 90 m
Swath Width	60 km	**Swath Width**	60 km	**Swath Width**	60 km
Quantization	8 bits	**Quantization**	8 bits	**Quantization**	12 bits

From Jensen (2007)

Table 2-8: Sensor characteristics of Space Imaging/GeoEye, Inc., *IKONOS* satellite, ORBIMAGE, Inc., *OrbView-3* Satellite, and DigitalGlobe, Inc., *QuickBird* satellite

Space Imaging/GeoEye, Inc.			ORBIMAGE, Inc.			DigitalGlobe, Inc.		
IKONOS			OrbView-3			QuickBird		
Band	Spectral Res. (μm)	Spatial Resolution (m) at Nadir	Band	Spectral Res. (μm)	Spatial Resolution (m) at Nadir	Band	Spectral Res. (μm)	Spatial Resolution (m) at Nadir
1	0.45 – 0.52	4 X 4	1	0.45 – 0.52	4 X 4	1	0.45 – 0.52	2.44 X 2.44
2	0.52 – 0.60	4 X 4	2	0.52 – 0.60	4 X 4	2	0.52 – 0.60	2.44 X 2.44
3	0.63 – 0.69	4 X 4	3	0.625 – 0.695	4 X 4	3	0.63 – 0.69	2.44 X 2.44
4	0.76 – 0.90	4 X 4	4	0.76 – 0.90	4 X 4	4	0.76 – 0.89	2.44 X 2.44
Pan	0.45 – 0.90	1 X 1	Pan	0.45 – 0.90	1 X 1	Pan	0.45 – 0.90	0.61 X 0.61
Sensor	Linear array pushbroom		Sensor	Linear array pushbroom		Sensor	Linear array pushbroom	
Swath	11 km		Swath	8 km		Swath	20 to 40 km	
Rate	25 Mb/s		Rate	50 Mb/s		Rate	50 Mb/s	
Revisit	< 3 days		Revisit	< 3 days		Revisit	1 to 5 days depend on latitude	
Orbit	681 km, Sun-synchronous, Equatorial crossing 10-11 a.m.		Orbit	470 km, Sun-synchronous, Equatorial crossing 10-11 a.m.		Orbit	600 km, not Sun-synchronous, Equatorial crossing variable	
Launch	April 27,1999 (failed), September 24,1999		Launch	June 26,2003		Launch	October 18,2001	

From Jensen (2007)

Table 2-9: Thailand Earth Observation Satellite (THEOS) sensor system characteristics.

Band	Spectral Res. (nm)	Spatial Res. (m) at nadir
Multispectral Band 1	450 – 520	15 X 15
Multispectral Band 2	530 – 600	15 X 15
Multispectral Band 3	620 – 690	15 X 15
Multispectral Band 4	770 – 900	15 X 15
Panchromatic Band	450 – 900	2.5 X 2.5
Sensor	Linear CCD array	
Swath	90 km+/- 30 Deg (Multispectral Band) 22 km+/- 30 Deg (Panchromatic Band)	
Rate	120 Mb/s	
Revisit	26 days	
Orbit	822 km, Sun-synchronous, Equatorial crossing 10.00 a.m.	

Table 2-10: Leica Geostsystems Airborne Digital Sensors 40 (ADS-40) characteristics.

Band	Spectral Resolution (nm)	Spatial Resolution (cm) at Nadir
Pan	465 – 680	Variable 15 cm @ 1450 m AGL
Blue	430 – 490	Variable
Green	535 – 585	Variable
Red	610 – 660	Variable
Near-infrared	835 – 885	Variable
Sensor	7 linear array CCDs	
Swath	12,000 pixels	
Detector size	6.5 µm	
Detector per linear array	Panchromatic = 2 x 12,000, Multispectral = 12,000	
Radiometric Resolution	12 bit	
Focal Length	62.7 mm	
Largest Application Scale	1:500 – 1:5,000 (depending upon aircraft altitude, AGL)	
Geo-referencing	Applanix POS IMU with GPS and INS	

From Jensen (2007)

Table 2-11: Characteristics of the NASA Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and the ITRES Research, Ltd., Compact Spectrographic Imager 3 (CASI 3) hyperspectral remote sensing systems.

Sensor	AVIRIS	CASI 3
Technology	Whiskbroom Linear Array	Linear (1480) and area array CCD (1480 X 288)
Spectral Res. (nm)	400 – 2500	400 – 1000
Spectral Interval (nm)	10	2.2
No. of Bands	224	228 (possible)
Quantization (bits)	12	14
IFOV (mrad)	1.0	0.49
Total FOV (Deg.)	30	40.5

From Jensen (2007)

Table 2-12: Characteristics of the Terra satellite Moderate Resolution Imaging Spectrometer (MODIS)

Band	Spectral Res.(μm)	Spatial Res.	Band Utilities
1	0.620-0.670	250X250m	Land-cover classification, chlorophyll absorption, leaf-area-index mapping
2	0.841-0.876	250X250m	
3	0.459-0.479	500X500m	Land, cloud, and aerosol properties
4	0.545-0.565	500X500m	
5	1.230-1.250	500X500m	
6	1.628-1.652	500X500m	
7	2.105-2.155	500X500m	
8	0.405-0.420	1X1Km	Ocean color, phytoplankton, biogeochemistry
9	0.438-0.448	1X1Km	
10	0.483-0.493	1X1Km	
11	0.526-0.536	1X1Km	
12	0.546-0.556	1X1Km	
13	0.662-0.672	1X1Km	
14	0.673-0.683	1X1Km	
15	0.743-0.753	1X1Km	
16	0.862-0.877	1X1Km	
17	0.890-0.920	1X1Km	Atmospheric water vapor
18	0.931-0.941	1X1Km	
19	0.915-0.965	1X1Km	

Table 2-12: (continued)

Band	Spectral Res.(μm)	Spatial Res.	Band Utilities
20	3.600-3.840	1X1Km	Surface-cloud temperature
21	3.929-3.989	1X1Km	
22	3.929-3.989	1X1Km	
23	4.020-4.080	1X1Km	
24	4.433-4.498	1X1Km	Atmospheric temperature
25	4.482-4.549	1X1Km	
26	1.360-1.390	1X1Km	Cirrus clouds
27	6.535-6.895	1X1Km	Water vapor
28	7.175-7.475	1X1Km	
29	8.400-8.700	1X1Km	
30	9.580-9.880	1X1Km	Ozone
31	10.780-11.280	1X1Km	Surface-cloud temperature
32	11.770-12.270	1X1Km	
33	13.185-13.485	1X1Km	Cloud-top altitude
34	13.485-13.785	1X1Km	
35	13.785-14.085	1X1Km	
36	14.085-14.385	1X1Km	

From Jensen (2007)

Table 2-13: Leica Geosystems Emerge Digital Sensor System (DSS) characteristics

Band	Nat. Color (nm)	NIR (nm)	Spatial Res. (m) at Nadir
1	400 – 500	-	Variable: 0.15 to 3 m
2	500 – 600	510 – 600	Variable: 0.15 to 3 m
3	600 – 580	600 – 700	Variable: 0.15 to 3 m
4	-	800 – 900	Variable: 0.15 to 3 m
Sensor	Area Array (4092 X 4079)		
Swath	4092 pixels		
Altitude	1 to 20000 ft.		
Pixel Size	0.009 mm		
Quantization	8 – and 16 bits		

From Jensen (2007)

Furthermore, Jensen (2007) summarized a selected list of some of the most important remote sensing systems and their characteristics as shown in Table 2-14.

Table 2-14: Selected remote sensing systems and their characteristics.

Remote Sensing Systems	Resolution: Spectral: Blue	Green	Red	Near-infrared	Middle-infrared (SWIR)	Thermal infrared	Micro-wave	Spatial (m)	Temporal (days)
Suborbital Sensors									
Panchromatic film (black & white)	0.5 — 0.7 μm							Variable	Variable
Color film	0.4 — 0.7 μm							Variable	Variable
Color-infrared film	0.5 — 0.9 μm							Variable	Variable
Digital Frame Cameras (CCD)	1	1	1	1				0.25 – 5	Variable
CASI - 1500	0.40 — 288 bands — 1.0 μm							variable	Variable
AVIRIS - Airborne Visible Infrared Imaging Spectrometer	0.40 — 224 bands — 2.5 μm							2.5 or 20	Variable
Intermap Star-3*i* X-band radar							1	Variable	Variable
Satellite Sensors									
NOAA-9 AVHRR LAC			1	1		3		1100	14.5/day
NOAA- K, L, M			1	1	2	2		1100	14.5/day
Landsat Multispectral Scanner (MSS)		1	1	2				79	16 - 18
Landsat 4 and 5 Thematic Mappers (TM)	1	1	1	1	2	1		30 and 120	16
Landsat 7 Enhanced TM (ETM+) — Multispectral	1	1	1	1	2	1		30 and 60	16
— Panchromatic	0.52 — 0.9 μm							15	16
SPOT 4 HRV — Multispectral		1	1	1				20	Pointable
— Panchromatic	0.51 — 0.73 μm							10	Pointable
GOES Series (East and West)	0.52 — 0.72 μm					4		700	0.5/hr
European Remote Sensing Satellite (ERS-1 and 2)	VV polarization C-band (5.3 GHz)						1	26 – 28	-
Canadian RADARSAT (several modes)	HH polarization C-band (5.3 GHz)						1	9 – 100	1 - 6 days
Shuttle Imaging Radar (SIR-C)							3	30	Variable
Sea-Viewing Wide Field-of-View Sensor (SeaWiFS)	3	2	1	2				1130	1
MODIS - Moderate Resolution Imaging Spectrometer	0.405 — 36 bands — 14.385 μm							250, 500, 1000	1 - 2
ASTER - Advanced Spaceborne Thermal Emission and Reflection Radiometer	0.52 — 3 bands — 0.86 μm							15	5
					1.6 — 6 bands — 2.43 μm			30	16
					8.12 — 5 bands — 11.6 μm			90	16
MISR - Multiangle Imaging SpectroRadiometer	Nine CCD cameras in four bands (440, 550, 670, 860 nm)							275 and 1100	1 - 2
NASA Topex/Poseidon — TOPEX radar altimeter	(18, 21, 37 GHz)							315,000	10
— POSEIDON single-frequency radiometer	(13.65 GHz)								
Space Imaging IKONOS — Multispectral	1	1	1	1				4	Pointable
— Panchromatic	0.45 — 0.9 μm							1	
Digital Globe QuickBird — Multispectral	1	1	1	1				2.44	Pointable
— Panchromatic	0.45 — 0.9 μm							0.61	

Form: Jensen (2007)

Chapter 3: Aerial Photography and Visual Interpretation

Aerial photograph was the first major remote sensing technique and even today it is still widely used. Jensen (2007) claimed that there are a number of important reasons why photo or image interpretation is such a powerful scientific tools, including:

- the aerial/regional perspectives;
- three-dimensional depth perception;
- knowledge beyond our human visual perception;
- the ability to obtain a historical image record to document change.

3.1 Characteristics of Aerial Photographs

Sabins (1987) stated that aerial photographs are acquired with a variety of cameras, films, and filters. Characteristics such as resolution, scale, and relief displacement are common to all photographs.

3.3.1 Resolution

In general, we consider resolution into two categories: spatial resolution and ground resolution.

A. Spatial resolution of photographs. Spatial resolution or resolving power of aerial photographs is influence by several factors:

1.1 Atmospheric scattering, which results from multiple interactions between light rays and the gases and particles of the atmosphere.

1.2 Vibration and motion of the aircraft, which are minimized by vibration-free camera mounts and motion compensation devices.

1.3 Resolving power of lenses. The resolving power of a lens is determined by its optical quality and size. If a lens is used to photograph a resolution target, such as the one shown in Figure 3-1, there is an upper limit to the number of line-pairs within the space of a millimeter that can be resolved on the resulting photograph. This maximum number of resolvable line-pairs per millimeter is a measure of the resolving power of the lens.

1.4 Resolving power of films. The resolving power of film is determined by several factors, the most important of which is the film's granularity. The two factors that largely determine granularity are the size distribution of silver halide grains in the emulsion and the nature of the development process. Films with high granularity have lower resolving power than those with low granularity. There is a trade-off between granularity and the speed of film: films with high granularity are faster, meaning they are more sensitive to light.

One method for expressing the resolving power of film is to photograph a resolution target and determine the maximum number of line-pairs per millimeter one can distinguish on the developed films as shown in Figure 3-2, targets with high contrast ratios produce better resolution than those with low contrast ratios; terrain features typically have low contrast ratios.

Film resolving power is commonly stated for both targets with high contrast ratios and those with low contrast ratios. For example, a widely used black-and -white film, Kodak Panotomic X aerial film, has a resolving power of 300 line-pairs mm-1 for high contrast target and 80 line-pairs mm-1 for low-contrast targets. System resolution (Rs) of a camera and film combination results from the resolving power off the lens and films and typically ranges from about 25 to 100 line-pairs mm-1.

B. Ground Resolution. Ground resolution expresses the ability to resolve ground features on aerial photographs. Systems resolution is converted into ground resolution by the formula:

$$R_g = \frac{H}{R_s \cdot F} \tag{3-1}$$

where:

R_g is ground resolution in line-pairs per millimeter

H is camera height above ground in meters

R_s is system resolution in line-pairs per millimeter

f is camera focal length in millimeters.

Figure 3-3 shows the geometric basis for this relationship.

Minimum ground separation is the minimum distance between two objects on the ground at which they can be resolved on the photograph. As Figures 3-3 shows, it is the separation between lines or bars in the resolution target and is determined by

$$Minimum\ ground\ separation = \frac{1.0\ line\text{-}pairs/R_g}{2} \tag{3-2}$$

Table 3-1 lists minimum ground separation values for typical aerial photographs acquired with camera system of medium and high resolution.

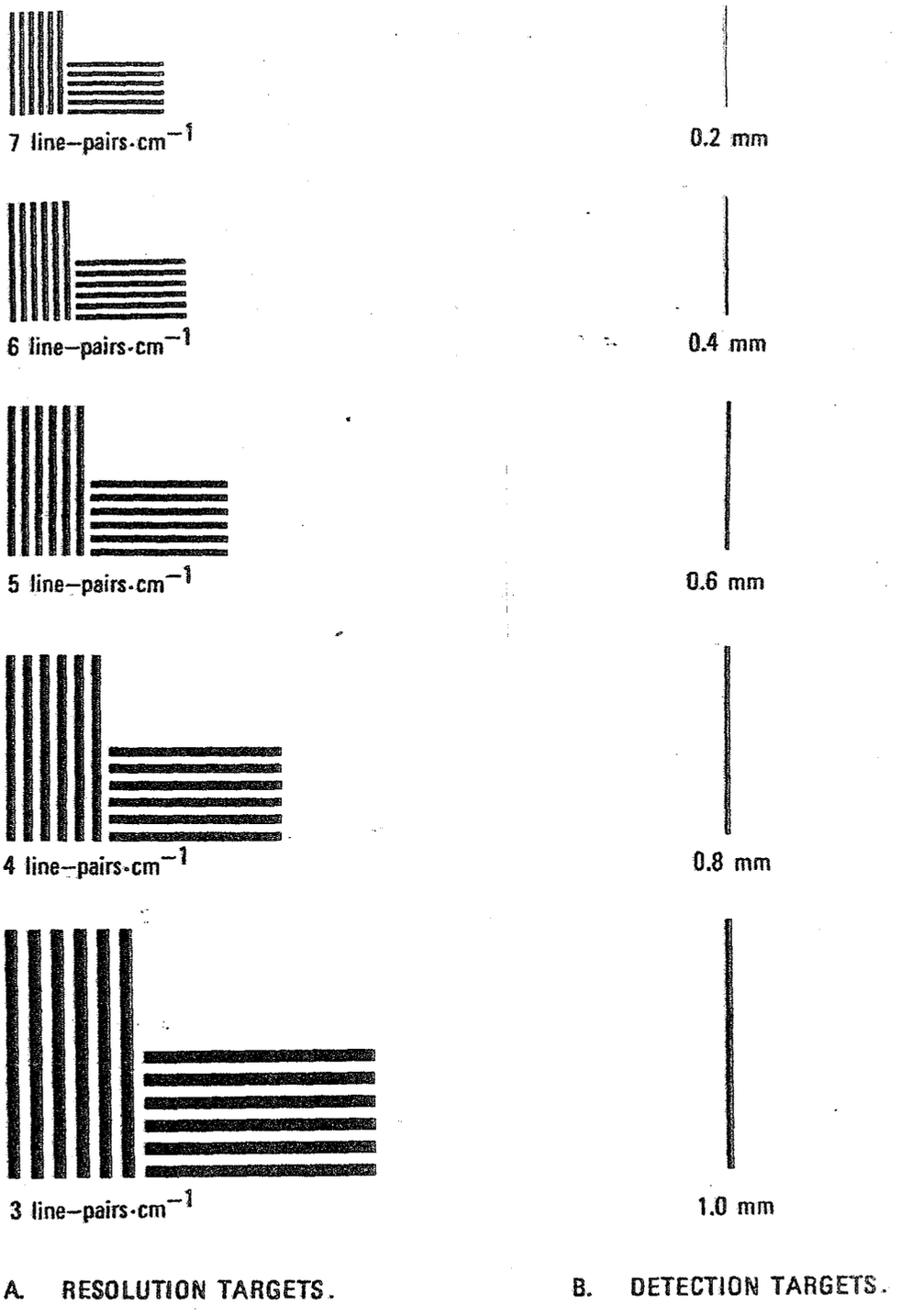


Figure 3-1: Resolution and detection targets with high contrast ratio. View chart form a distance of 5 m.For A, determine the most closely spaced set of bars you can resolve. For B, determine the narrowest bar you can detect (From Sabins, 1987)

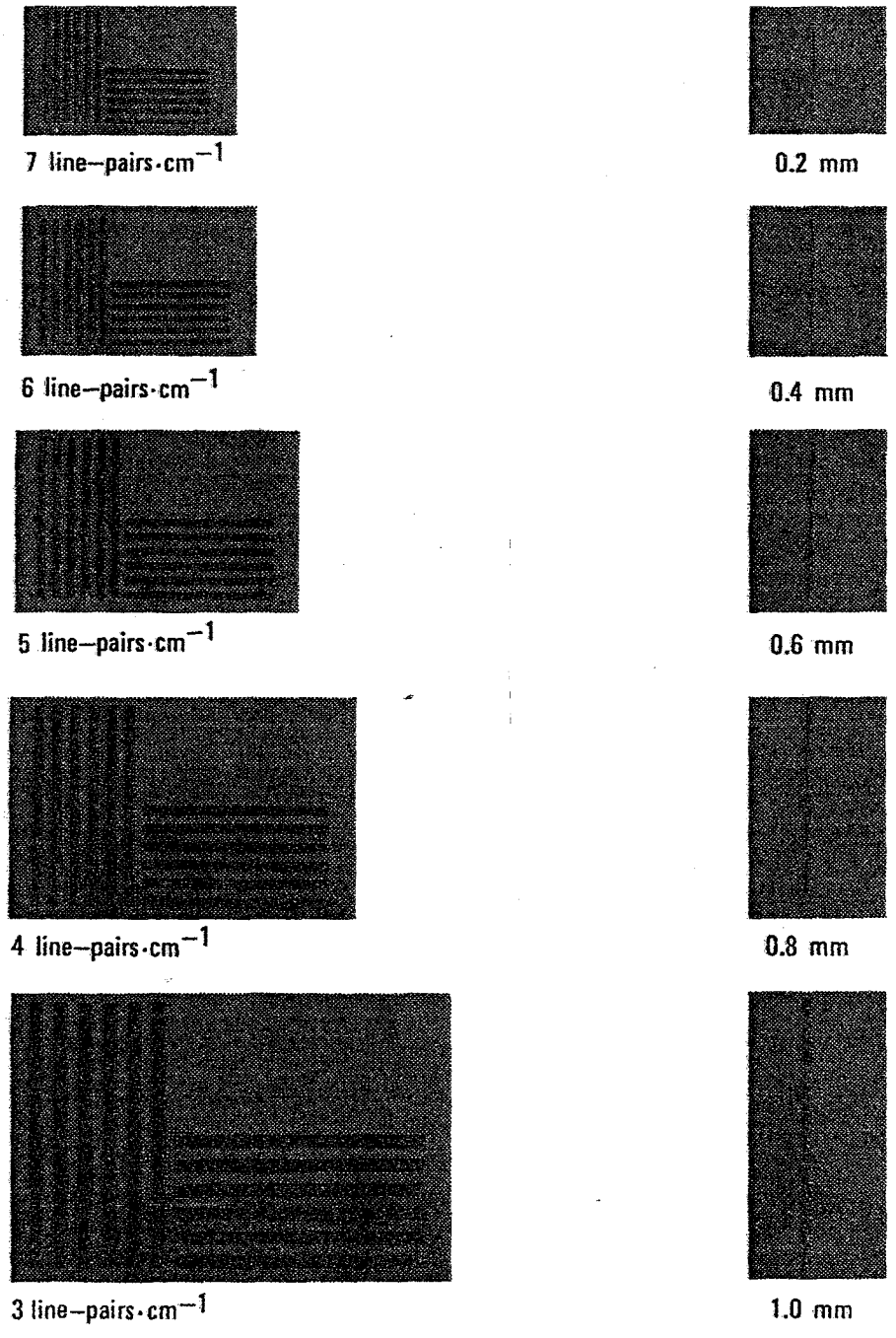


Figure 3-2: Resolution and detection targets with low contrast ratio. View chart form a distance of 5 m. For A, determine the most closely spaced set of bars you can resolve. For B, determine the narrowest bar you can detect (From Sabins, 1987)

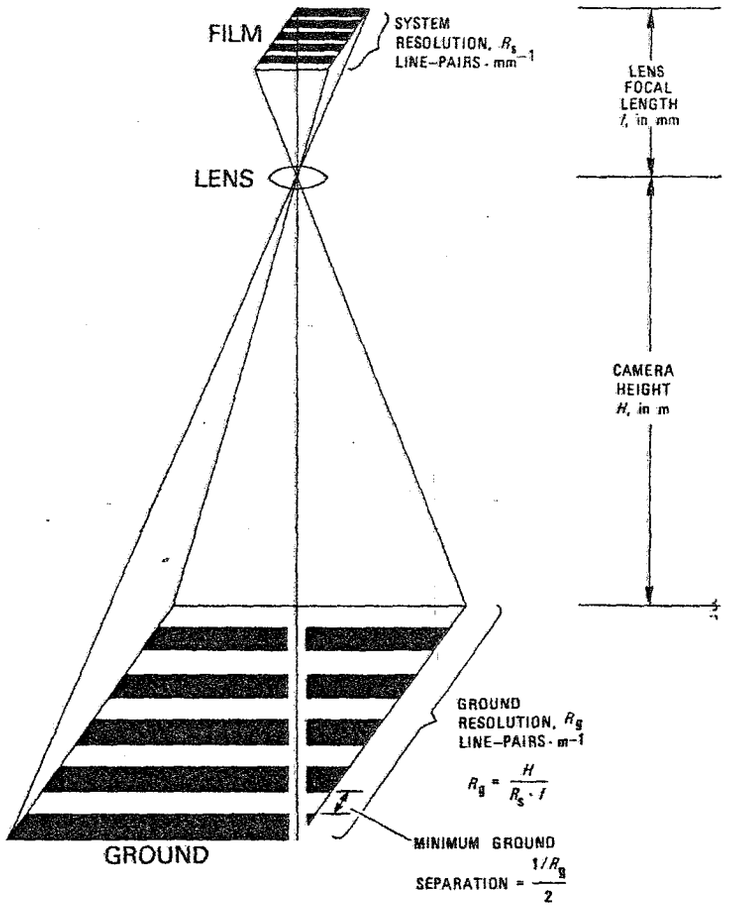


Figure 3-3: Ground resolution and minimum ground resolution on aerial photographs (From Sabins, 1987).

Table 3-1: Minimum ground separation on typical aerial photographs acquired at different height (focal length of camera lens is 152 mm)

Camera height (H), m	Scale of photographs	Minimum ground separation for system resolution (Rs) of	
		40 line-pairs mm^{-1}, m	100 line-pairs mm^{-1},m
1,525	1 : 10,000	0.12	0.05
3,050	1 : 20,000	0.25	0.10
4,575	1 : 30,000	0.37	0.15
6,100	1 : 40,000	0.50	0.20

From: Sabins, 1987.

3.3.2 Photographic Scale.

Scale of aerial photographs is determined by the relationship:

$$Scale = \frac{1}{H/f} \tag{3-3}$$

where:

H is camera height above ground

f is camera focal length

Both H and f must be given in the same units, typically meters.

3.3.3 Relief Displacement.

Relief displacement is present on all vertical aerial photographs that are acquired with the camera aimed directly down. The tops of objects such as building appear to "lean" away from the principal point, or optical center, of the photograph. The amount of displacement increases at greater radial distances from the center and reaches a maximum at the corners of the photograph. Figure 3-4 shows the geometry of image displacement, where lights ray are traced from the terrain through the camera lens and on to the film. Prints made from the film appear as though they were in the position shown by the plane of photographic print in Figure 3-4(a). The vertical arrows on the terrain represent objects of various heights located at various distances from the principal point. The light ray reflected from the base of object A intersects the plane of the photographic print at position A, and the ray from the top (or point of the arrow) intersects the print at A'. The distance A-A' is the relief displacement (d) shown in the plane view (Figure 3-4(b).

The amount of relief displacement (d) on an aerial photograph is

1. directly proportional to the height (h) of the object. For objects A and C (Figure 3-4) at equal distances from the principal point, d is greater for A, which is the taller object.

2. directly proportional to the radial distance (r) from the principal point on the displacement image corresponding to the object (Figure 3-4) For objects A and B, which are of equal height, d is greater for A because further from the principal point.

3. inverse proportional to the height (H) of the camera above the terrain.

These relations are expressed mathematically as

$$d = \frac{h \cdot r}{H} \tag{3-4}$$

which may be transposed to

$$h = \frac{H \cdot d}{r} \tag{3-5}$$

where:

h is the height of object

H is the height of the camera

d is relief displacement

r is radial distance from the principal point to the point on the displaced image corresponding to the top of the object.

This equation may be used to determine the height of an object from its relief displacement on an aerial photograph.

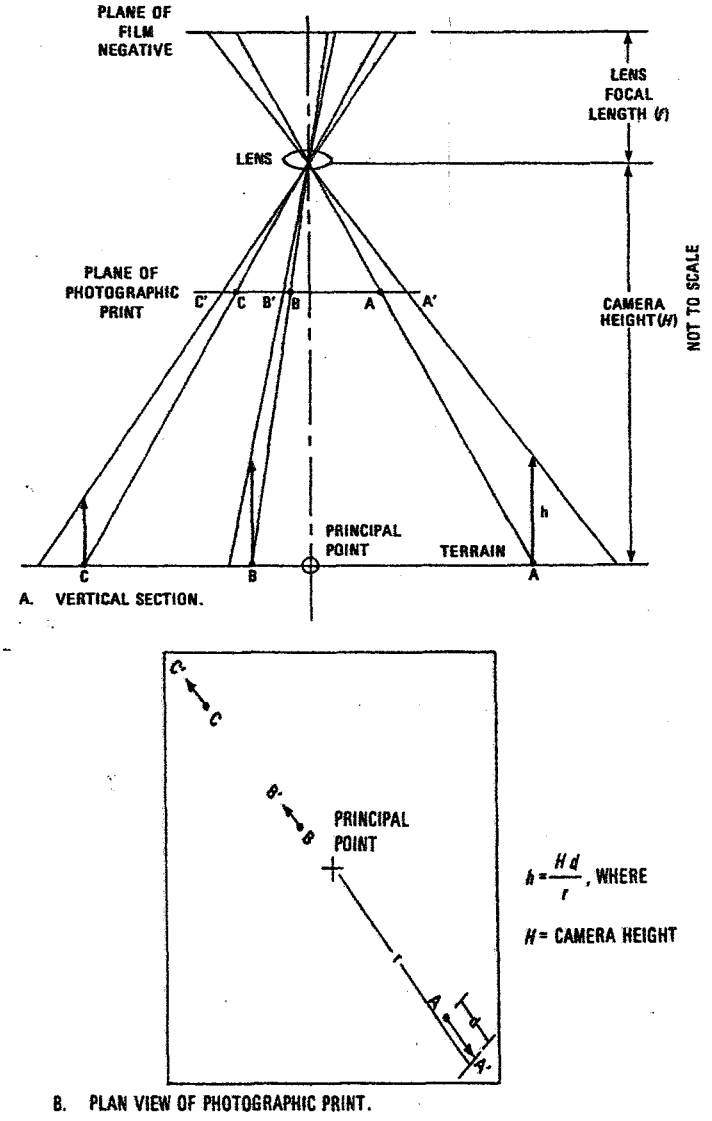


Figure 3-4: Geometry of relief displacement on a vertical aerial photograph (From Sabins, 1987).

3.2 Visual Interpretation

Two main interpretation procedures in remote sensing are visual interpretation and digital image analysis. Estes et al. (1983) suggested a broad conceptual framework of image analysis that included methods for detecting, identifying, and measuring objects of interest from the aerial perspective. Both manual and computer-assisted procedures of remotely sensed data have the same general goals and share a basic element of image interpretation.

Jensen (2007) stated both analog and digital analysis of remotely sensed data seek to detect and identify important phenomena in the scene. Once identified, the phenomena are usually measured, and the information is used in solving problems. Optimum results are often achieved using a synergistic combination of both visual and digital image processing.

Campbell (1987) stated that proficiency in image interpretation come from three kinds of knowledge: subject, geographic region, and remote sensing system.

A. Subject: Knowledge of the subject of interpretation, which is the kind of information to examine from the image, must form the heart of interpretation. Accurate interpretation requires familiarity with the subject of the interpretation. For example, interpretation of geologic information requires education and experience in the field of geology. Image interpreters should be equipped with a broad range of knowledge pertaining to the subject at hand and their interrelationships.

B. Geographic region: Knowledge of the specific geographic region depicted on an image can be equally significant. In unfamiliar regions the interpreter may find it necessary to make a field reconnaissance or to use maps and books that describe analogous regions with similar climate, topography, or land use.

C. Remote sensing system: Knowledge of the remote sensing system is obviously essential. The interpreter must understand how each image is formed and how each sensor portrays landscape features. Different instruments use separate portions of the electromagnetic spectrum, operate at different resolutions, and use different methods of recording images. The image interpreter must know how each of these variables influences the image to be interpreted and how to evaluate their effects on his ability to derive useful information from the imagery.

3.2.1 Image Interpretation Task

The image interpreter must routinely several kinds of tasks, many of which may be completed together in an integrated process. Four main image interpretations are classification, enumeration, mensuration, and delineation (Campbell, 1987)

A. Classification

The first task is classification, assigning objects, features, or areas to classes based upon their appearance on the imagery. Often the distinction is made between levels of confidence and precision. ***Detection*** is the determination of presence or absence of a feature. ***Recognition*** implies a higher level of knowledge about a feature or object, such that the object can be assigned an identity in a general class or category. Finally ***identification*** means that the identity of an object or feature can be specified with enough confidence and detail to place it in a very specific class.

B. Enumeration

Enumeration refers to listing and counting discrete items on image. For example, housing units can be classified as "detached single family," "multifamily complex;" "mobilehome," "multiresidential," and so on, and then reported as number present within a defined area. Clearly the ability to conduct such an enumeration depends upon to classify items accurately as discussed above.

C. Mensuration

Measurement, or mensuration, is an importance function in many interpretation problems. Two kinds of measurement are important. First, is the measurement of distance and height, and by extension, of volumes and areas as well. The practice of making such measurements forms the subject of *photogrammetry*, which applies a knowledge of image geometry to derivation of accurate distance. A second form of measurement is quantitative assessment of image brightness. The science of *photometry* is devoted to measurement of the intensity of light, including estimation of scene brightness by examination of image tone, using a special instrument known as a densitometer. If the measured radiation extends outside the visible spectrum, the term *radiometry* applies. Photometry and radiometry apply similar instruments and principles, so the are closely related to one another.

D. Delineation

Finally, the interpreter must often delineate, or outline, regions as they are observed on remotely sensed images. The interpreter must be able to separate distinct areal units that are characterized by specific tones and textures and to identify edges, or boundaries, between separate areas. Typical examples include delineation of separate class of forest or land use.

3.2.2 Element of image interpretation

By tradition, image interpreters are said to employ some combination of eight element of image interpretation, which describe the characteristics of objects and features as they appear on remotely sensed image (Campbell, 1987)

A. Image Tone: Image tone denoted the lightness or darkness of a region within an image. For black-and-white images, tone may be characterized as "light," "medium gray," "dark gray," "dark", and so on, as the image assumed varied shaded of white, gray, or black. For color or color infrared imagery, image tone refers simply to "color" described informally perhaps in such term as "dark green", "light blue", or "pale pink". Image tone refers ultimately to the brightness of an area of ground as portrayed by the film in a given spectral region or in three spectral regions for color or color infrared film (Figure 3-5).

B. Image Texture: Image texture refers to the apparent roughness or smoothness of an image region. Usually texture is caused by the pattern of highlighted and shadowed areas as an irregular surface is illuminated from an oblique angle (Figure 3-6).

C. Shadow: Shadow is an especially important clue in the interpretation of objects. Shadow has been of great significance in distinguishing subtle difference that might not otherwise by visible. By extension, the role of shadow in interpretation of any made landscape in which kinds of structures or subjects be identified is significant (Figure 3-7).

D. Pattern: Pattern refers to the arrangement of individual objects into distinctive, recurring forms that permit recognition on imagery. Pattern on an image usually follows a functional relationship between the individual features that compose the pattern. For example, the buildings in an industrial plant may have a distinctive pattern because they are organized to permit economical flow of material through the plant from receiving raw material to shipping the finished product (Figure 3-8).

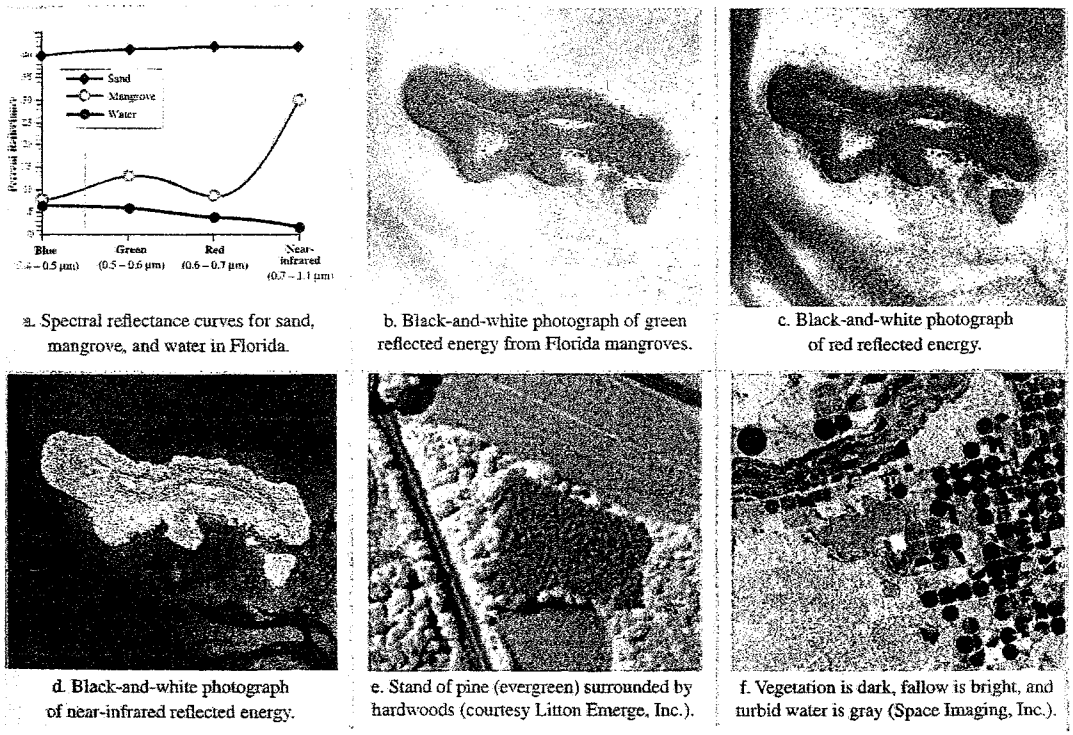
E. Association: Association specified characteristics occurrence of certain objects or features, usually without the strict spatial arrangement implied by pattern (Figure 3-9).

F. Shape: Shape of features are obvious clues to their identities. For example, individual structure and vehicles have characteristics shapes, which, if visible in sufficient detail, provide the basis for identification. Features in nature often have such distinctive shapes that shape alone might be sufficient to provide clear identification (Figure 3-10).

G. Size: Size is important in two ways. First is the relative size of an object or feature in relation to other objects on the image. The role is achieved by recognition of familiar objects (dwellings, highways, river, etc.), then extrapolation to relate sizes of these known features to estimate sizes and identities of those might not be easily identified. Secondly, absolute measurement can be equally valuable as an interpretation aid. Measurement of the size of an object can confirm its identification based upon other factors, especially if its dimensions are so distinctive that form definitive criteria for specific items or classes of items (Figure 3-11).

H. Site: Site refers to topographic position. For example, sewage treatment facilities are positioned at low topographic site, near streams or rivers, to collect waste flowing through the system from higher locations (See also Figure 3-9).

Jensen (2007) summarized the qualitative and quantitative characteristics of elements of image interpretation as shown in Table 3-2.



a. Spectral reflectance curves for sand, mangrove, and water in Florida.

b. Black-and-white photograph of green reflected energy from Florida mangroves.

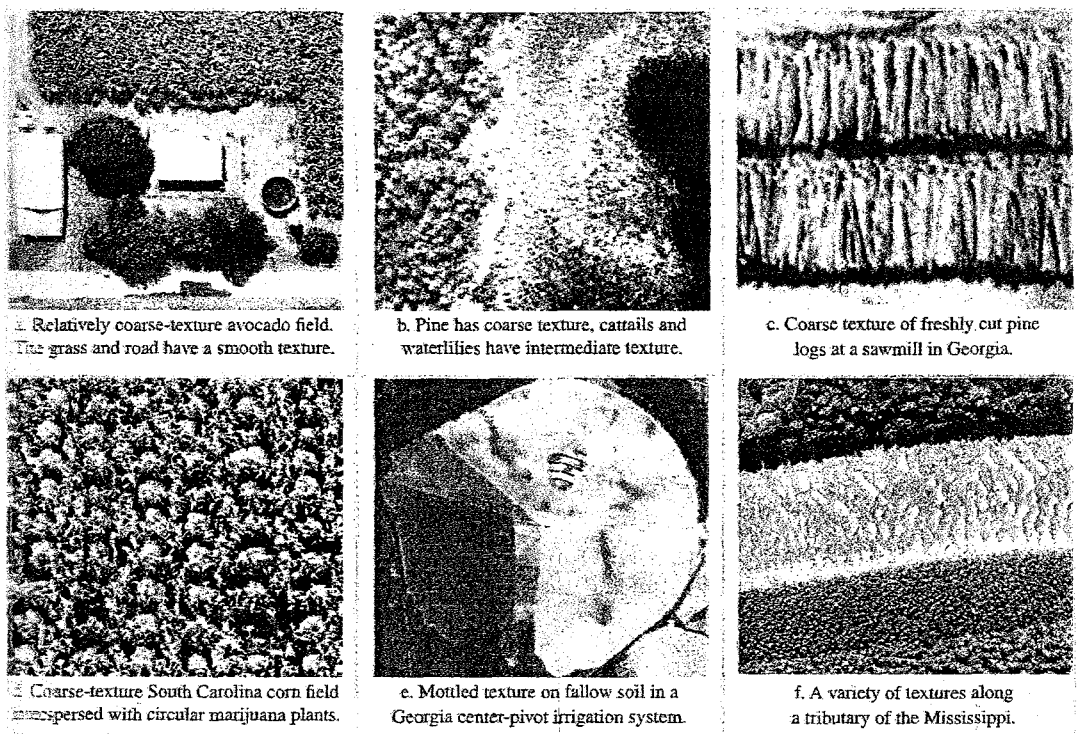
c. Black-and-white photograph of red reflected energy.

d. Black-and-white photograph of near-infrared reflected energy.

e. Stand of pine (evergreen) surrounded by hardwoods (courtesy Litton Emerge, Inc.).

f. Vegetation is dark, fallow is bright, and turbid water is gray (Space Imaging, Inc.).

Figure 3-5: Element of Image Interpretation - Tone (From Jensen, 2007)



a. Relatively coarse-texture avocado field. The grass and road have a smooth texture.

b. Pine has coarse texture, cattails and waterlilies have intermediate texture.

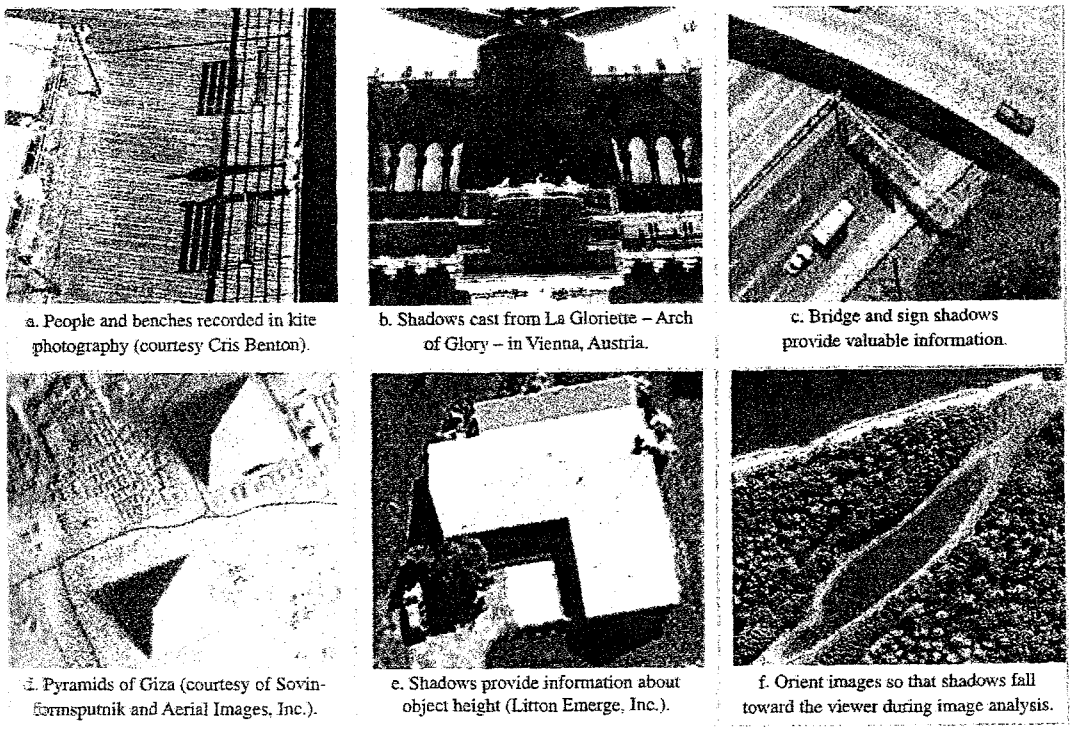
c. Coarse texture of freshly cut pine logs at a sawmill in Georgia.

d. Coarse-texture South Carolina corn field interspersed with circular marijuana plants.

e. Mottled texture on fallow soil in a Georgia center-pivot irrigation system.

f. A variety of textures along a tributary of the Mississippi.

Figure 3-6: Element of Image Interpretation - Texture (From Jensen, 2007).



a. People and benches recorded in kite photography (courtesy Cris Benton).

b. Shadows cast from La Gloriette – Arch of Glory – in Vienna, Austria.

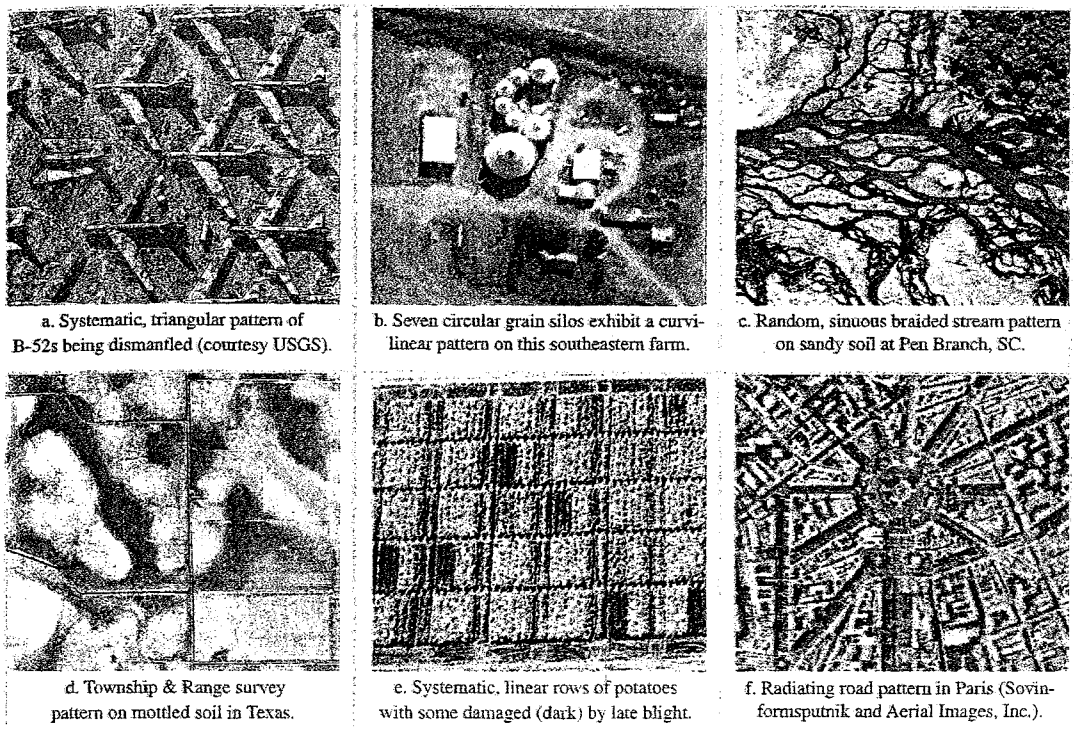
c. Bridge and sign shadows provide valuable information.

d. Pyramids of Giza (courtesy of Sovinformsputnik and Aerial Images, Inc.).

e. Shadows provide information about object height (Litton Emerge, Inc.).

f. Orient images so that shadows fall toward the viewer during image analysis.

Figure 3-7: Element of Image Interpretation - Shadow (From Jensen, 2007).



a. Systematic, triangular pattern of B-52s being dismantled (courtesy USGS).

b. Seven circular grain silos exhibit a curvilinear pattern on this southeastern farm.

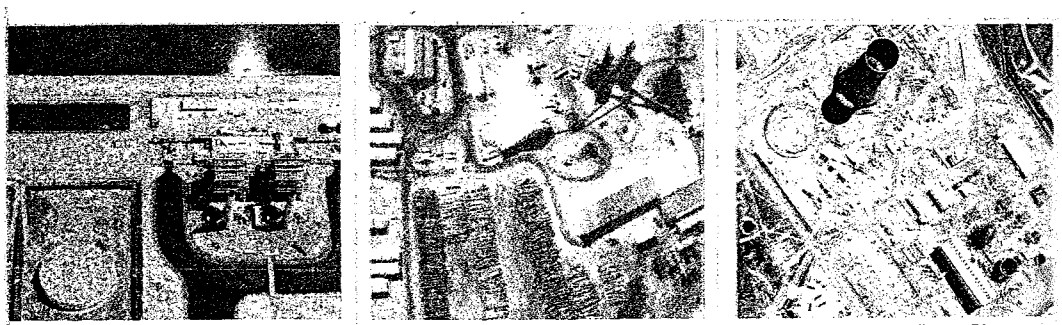
c. Random, sinuous braided stream pattern on sandy soil at Pen Branch, SC.

d. Township & Range survey pattern on mottled soil in Texas.

e. Systematic, linear rows of potatoes with some damaged (dark) by late blight.

f. Radiating road pattern in Paris (Sovinformsputnik and Aerial Images, Inc.).

Figure 3-8: Element of Image Interpretation - Texture (From Jensen, 2007).

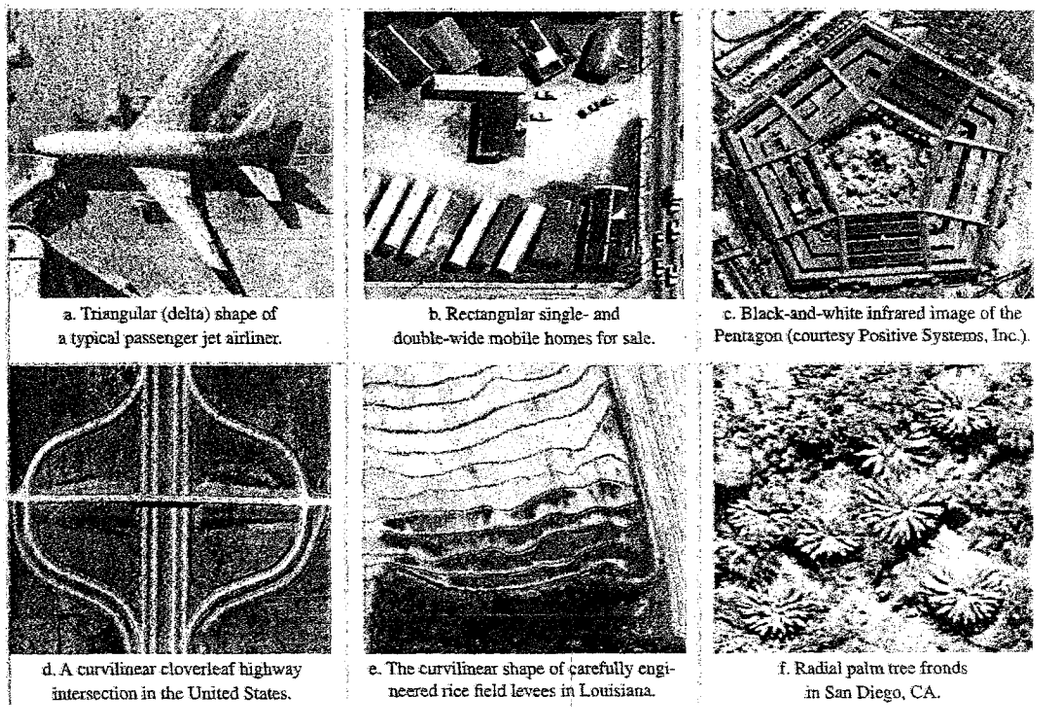


a. Thermal electric Haynes Steam Plant in Long Beach, CA.

b. A sawmill with its associated piles of raw and finished lumber.

c. Vogtle Nuclear Power Plant near Augusta, GA.

Figure 3-9: Element of Image Interpretation – Site, Situation and Association (From Jensen, 2007).



a. Triangular (delta) shape of a typical passenger jet airliner.

b. Rectangular single- and double-wide mobile homes for sale.

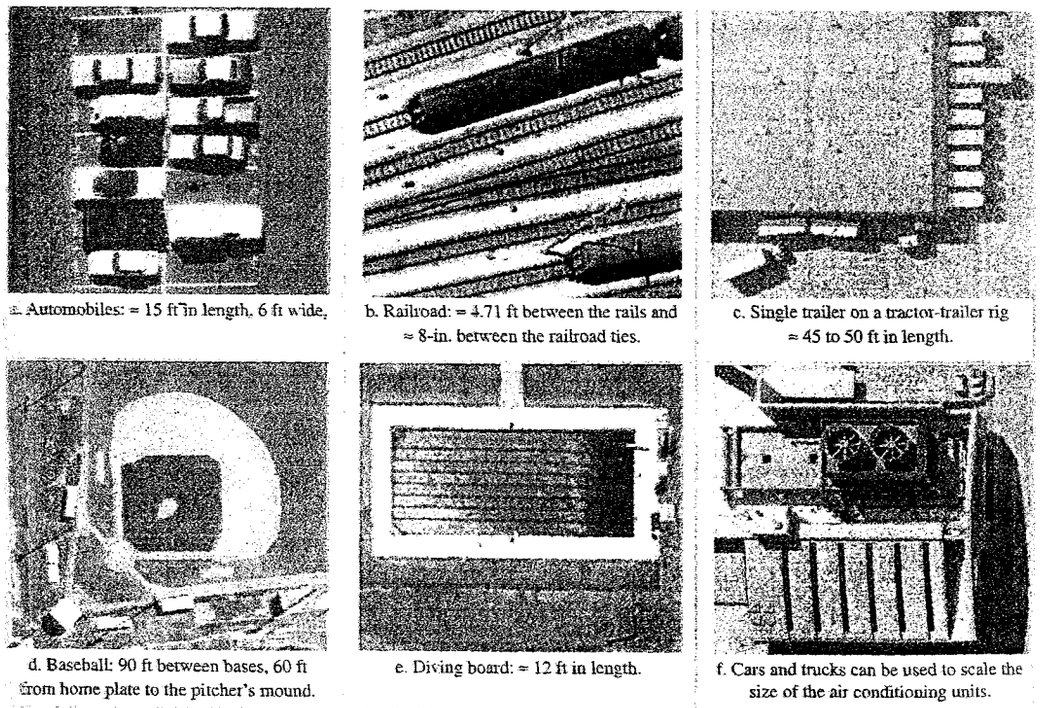
c. Black-and-white infrared image of the Pentagon (courtesy Positive Systems, Inc.).

d. A curvilinear cloverleaf highway intersection in the United States.

e. The curvilinear shape of carefully engineered rice field levees in Louisiana.

f. Radial palm tree fronds in San Diego, CA.

Figure 3-10: Element of Image Interpretation - Shape (From Jensen, 2007).



a. Automobiles: ≈ 15 ft in length, 6 ft wide.

b. Railroad: ≈ 4.71 ft between the rails and ≈ 8-in. between the railroad ties.

c. Single trailer on a tractor-trailer rig ≈ 45 to 50 ft in length.

d. Baseball: 90 ft between bases, 60 ft from home plate to the pitcher's mound.

e. Diving board: ≈ 12 ft in length.

f. Cars and trucks can be used to scale the size of the air conditioning units.

Figure 3-11: Element of Image Interpretation - Size (From Jensen, 2007).

Table 3-2: Element of image interpretation

Element	Common Adjectives (quantitative and qualitative)
x, y Location	• x,y coordinate; longitude and latitude or meters easting and northing in a UTM map grid
Size	• length, width, perimeters, area • small, medium (intermediate), large
Shape	• an object's geometric characteristics: linear, curvilinear, circular, elliptical, radial, square, rectangular, triangular. Hexagonal, pentagonal, star, amorphous, etc.
Shadow	• A silhouette caused by solar illumination from the side
Tone/color	• gray tone; light (bright), intermediate (gray), dark (Black) • Color: HIS=intensity, hue (color), saturation; RGB=red, green and blue; Munsell
Texture	• characteristic placement and arrangement of repetitions of tone or color: smooth, intermediate (medium), rough (coarse), mottled, stripped.
Pattern	• the spatial arrangement of objects on the ground; systematic, unsystematic or random, linear, curvilinear, rectangular, circular, elliptical, parallel, centripetal, serrated, striated, braided.
Height/depth volume/slope/Aspect	• z-elevation (height), depth (bathymetry), volume (m3), slope°, aspect°
Site/situation/ Association	• Site: elevation, slope, aspect, exposure, adjacent to water, transportation, utilities. • Situation: objects are placed in a particular order or orientation relative to one another • Association: related phenomena are usually present

From Jensen (2007)

3.2.3 Image Interpretation Strategies

An image interpretation strategy can be defined as a disciplined procedure that enables the interpretation to relate geographic pattern on the ground to their appearance on the image. Campbell (1978) defined five categories of image interpretation.

A. Field Observation: Field observation, as an approach to image interpretation, are required when the image and its relationship to ground conditions are so imperfectly understood that the interpreter is forced to go to the field to make an identification. In effect the analyst is unable to interpret the image from knowledge and experience at hard and must gather field observations to ascertain the relationship between landscape and its appearance on the image. Field observations are, of course, a routine dimension of any interpretation as a check on accuracy or a means of familiarization with a specific region. Mere their use as an interpretation strategy refers to the fact that when they are required for the interpretation, their use reflects a rudimentary understanding of the manner in which a landscape is depicted on a specific image.

B. Direct Recognition: Direct recognition is the application of an interpreter's experience, skill, and judgment to associate the image patterns with informational classes. This process is essentially a qualitative, subjective analysis of the image using the elements of the image interpretation as visual and logical clues. In everyday experience direct recognition is applied in an intuitive manner, for image analysis, it must be a disciplined process, with vary careful systematic examination of the image.

C. Interpretation by Inference: Interpretation by inference is the use of a variable distribution to map one that is not itself visible on the image. The visible distribution acts as a surrogate, or proxy (a substitute) for the mapped distribution. An example: Soils are defined by vertical profiles that cannot be directly observed by remotely sensed imagery. But soil distributions are sometime very closely related to patterns of landforms and vegetation that are recorded on the image. Thus, landforms and vegetation can form surrogates for the soil pattern; the interpreter infers the invisible soil distribution from patterns that are visible. Application of this strategy requires a complete knowledge of the link between the proxy and the mapped distribution; attempts to apply imperfectly defined proxies produce inaccurate interpretations.

D. Probabilistic Interpretation: Probabilistic interpretations are effort to narrow the range of possible interpretations by formally integrating non-image information into classification process, often by means of quantitative classification algorithms. For example, knowledge of the crop calendar can restrict the likely choices for identifying crops of a specific region. If it is known that winter wheat is harvested in June, the choice of crops for interpretation of an August image can be

restricted to eliminate wheat as a likely choice and thereby avoid a potential classification error. Often such knowledge can be expressed as a statement of probability. Possibly certain classes might favor specific topographic sites but occur over a range of sites, so a decision rule might express this knowledge as a 0.9 probability of finding the class on a well-drained site, but only a 0.05 probability of finding it on a poorly drained site. Several such statements systematically incorporated into the decision-making process can improve classification accuracy.

E. Deterministic Interpretation: A fifth strategy, deterministic interpretation, is the most rigorous and precise approach. Deterministic interpretations are based upon quantitatively expressed relationships that tie image characteristics to ground conditions. In contrast with the other methods most information is derived from the image itself. Photogrammetric analysis of stereo pairs for terrain information is a good example. A scene is imaged from two separate positions along a flight path, and the photogrammetrist measures the apparent displacement. Based upon knowledge of the geometry of the photographic system, a photographic model of the landscape can be reconstructed. The result is therefore the derivation of precise information about the landscape using only the image itself and a knowledge of its geometric relationship with the landscape. Relation to other methods, very little nonimage information is required.

3.2.4 Method of Search

Jensen (2007) suggested method of search to interpreting remotely sensed data including: (1) utilizing collateral (ancillary) information, (2) converging the evidence, and (3) applying the multi-concept in image analysis

A. Using Collateral Information: Trained image interpreters rarely interpret aerial photograph or other remote sensor data in a vacuum. Instead, they collect as much collateral (ancillary) information about the subject and the study area as possible. Some of the major types of collateral information are the use of a variety of maps for orientation, political boundary information, property line cadastral data, geodetic control (x, y, z), forest stand data, geologic data, hazard data, socioeconomic data, soil taxonomy, topographic and bathymetric data, transportation features, and wetland information. Ideally, these data are stored in a geographic information system (GIS) for easy retrieval and overlay with the remote sensor data.

B. Convergence of Evidence: *It is generally a good idea to work from the known to the unknown.* Careful examination of what we do know about things surrounding and influencing the object of interest can provide valuable clues that could allow us to make the identification. We bring all the knowledge we can to the interpretation problem and converge our evidence to identify the object or process at work.

C. The Multi-Concept: Jensen (2007) claimed that Robert Colwell of the Forestry Department at the University of California at Berkley put fourth the *multi-concept* in image interpretation in the 1960s. He suggested that the most useful and accurate method of scientific image interpretation consisted of performing the following types of analysis: *multispectral*, *multi-disciplinary*, *multi-scale*, and *multi-temporal*.

Colwell pioneered the use of *multi-band* aerial photography and *multi-spectral* remote sensor data. He documented that in agriculture and forest environment, measurements made in multiple discrete wavelength regions (bands) of the electromagnetic spectrum were usually more valuable than acquired single broad-band panchromatic-type imagery. He also suggested that *multi-scale* (often called *muti-stage*) photography or imagery of an area was very important. Smaller scale imagery (e.g., 1:80,000) was useful for placing intermediate scale imagery (e.g., 1:40,000) in its proper regional context. Then, very large scale imagery (e.g., > 1:10,000) could be used to provide detailed information about local phenomena. In situ field investigation is the largest scale utilized and is very important. Each scale of imagery provides unique information that can be used to calibrate the others.

Professor Colwell was a great believer in bringing many *multi-disciplinary* experts together to focus on a remote sensing image analysis or information extraction problem. He suggested that image analysis embrace the input of other multidisciplinary scientists in the interpretation process. This philosophy and process often yields synergistic, novel, and unexpected results. Colwell pioneered the concept of developing crop phenological calendars in order to monitor the spectral changes that take place as plant progress through the growing season. Once crop calendars are available, they may be used to select the optimum dates during the growing season to acquire remote sensor data. Table 3-3 lists the discipline of colleagues that often collaborate when systematically studying a certain topic.

In conclusion, image interpreters, of course, may apply a mixture of several strategies in a given situation. Interpretation of soil patterns, for example, may require direct recognition to identify specific classes of vegetation, then application of interpretation by proxy to relate the vegetative pattern to the underlying soil pattern.

Table 3-3: Multidisciplinary scientist brings their unique training to the image interpretation process.

Topic	Discipline
Agriculture	Agronomy, agricultural engineering, biology, biogeography, geology, landscape ecology, marine science
Database and algorithm preparation	Cartography, GIS, computer science, photogrammetry, programming, analytical modeling
Forestry, Rangeland	Forestry, agronomy, rangeland ecology, landscape ecology, biogeography
Geodetic control	Geodesy, surveying, photogrammetry
Geology, Soil	Geology, geomorphology, agronomy, geography
Hazard	Geology, hydrology, urban and physical geography
Hydrology	Hydrology, chemistry, geology, geography
Topography/Bathymetry	Geodesy, surveying, photogrammetry
Transportation	Transportation engineering, city planning, urban geography
Urban studies	Urban, economic, and political geography, city planning, transportation engineering, civil engineering, landscape ecology
Weather/ Atmosphere	Meteorology, climatology, physic, chemistry
Wetland	Biology, landscape ecology, biogeography

From Jensen (2007)

Chapter 4: Satellite Data and Digital Image Processing

4.1 Characteristics of Satellite Data

Digital images consist of discrete picture element, called pixels. Associated with each pixel is a number that is the average radiance (brightness) of a relatively small area within a scene. (Figure 4-1) The size of this area affects the reproduction of detail within the scene. As the pixel area is reduced, more scene detail is preserved in the digital representation. Digital images can be displayed at any desired scale by appropriate computer processing (Schowengerdt, 1983).

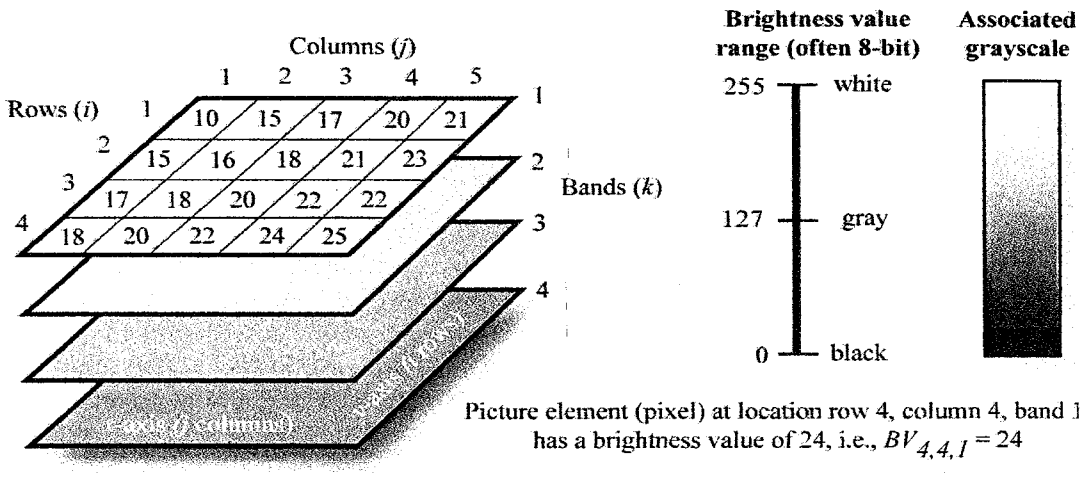


Figure 4-1: Digital remote sensor data (Jensen, 2005)

Bernstein (1983) defined a digital image as a sampled and quantized numeric representation of a scene. The scene is spatially partitioned by the sensing device into a regular array of numbers whose values represent the radiance or brightness of the sample region in one or more spectral bands (Figure 4-2). In addition to the spatial and brightness dimensions, a digital data set can have spectral and temporal dimensions since the scene can be viewed in many spectral bands, and at different times. Thus a digital image is a multidimensional matrix of numbers that characterize a scene.

A multispectral digital image (I) consisting of K spectral bands can then be represented as K arrays of M x N elements of non-negative value, or

$$I^{(k)} = (P_{ij}^{(k)}) \qquad (4\text{-}1)$$

where:

k is band number = 1, 2, ..., K,

i is line (row) number = 1, 2, ..., N,

j is element (sample or column) number = 1, 2, ..., M.

Each picture element (pixel) consists of B bits per sample. The total data element in a multispectral image is then **MNKB** bytes.

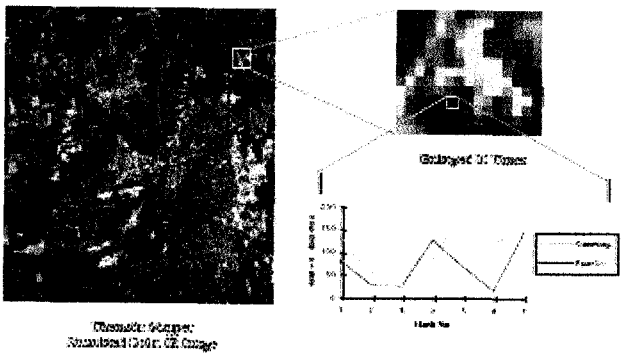


Figure 4-2 Multispectral concept and data representation (From Landgrebe, 2003)

Actually, digital numbers or brightness values from remotely sensed data are simply positive integers that result from quantizing the original electrical signal from the sensor into positive integer values using a process called analog-to-digital (A-to-D) signal conversion (Lillesand et al., 2004). Figure 4-3 is a graphical representation of the A-to-D conversion process. The electrical signal from the sensor is a continuous analog signal. This continuous signal is sampled at a set time interval (ΔT) and recorded numerically at each sample point (a,b, ... i, k). The sampling rate for a particular signal is determined by the highest frequency of change in the signal. The sampling rate must be at least twice as high as the highest frequency present in the original signal in order to adequately represent the variation in the signal. In this figure, the incoming sensor signal in terms of an electrical voltage value ranging between 0 and 2V. The DN output values are integers ranging from 0 to 255. Accordingly, a samples voltage of 0.46 recorded by the sensor at (a) would be recorded as a DN of 59.

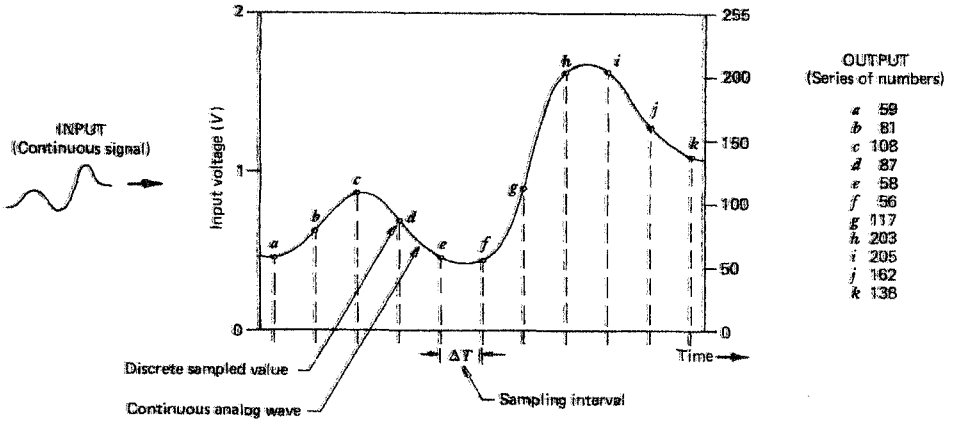


Figure 4-3: Analog-to-digital conversion process (From Lillesand et al., 2004).

Conversely, in an analog image, the brightness value is actually a surrogate for the density (D) of the light-absorbing silver or dye deposited at a specific location. The density characteristic of a negative or positive transparency film is measured using a densitometer (Jensen, 2005). There are several types of the densitometers including:

- Flatbed and drum microdensitometer (Figure 4-4 and Figure 4-5),
- Video densitometer (Figure 4-6),
- Linear or area charge-coupled device densitometer (Figure 4-7).

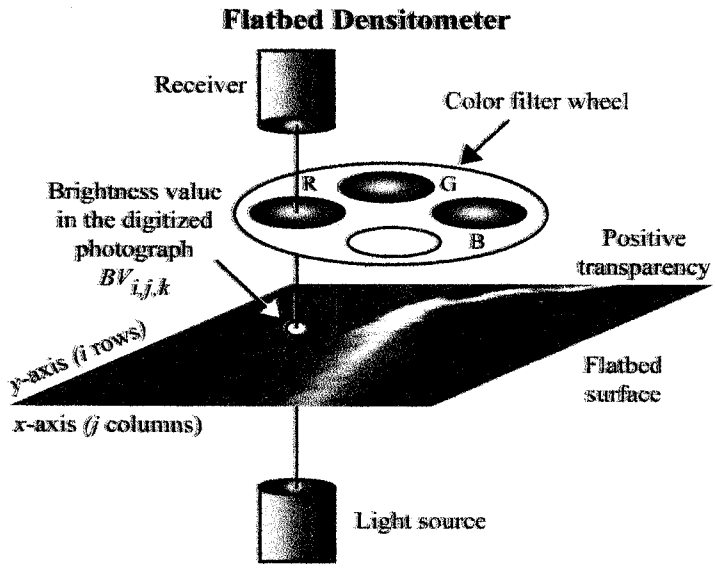


Figure 4-4: Schematic of flatbed densitometer (From Jensen, 2005).

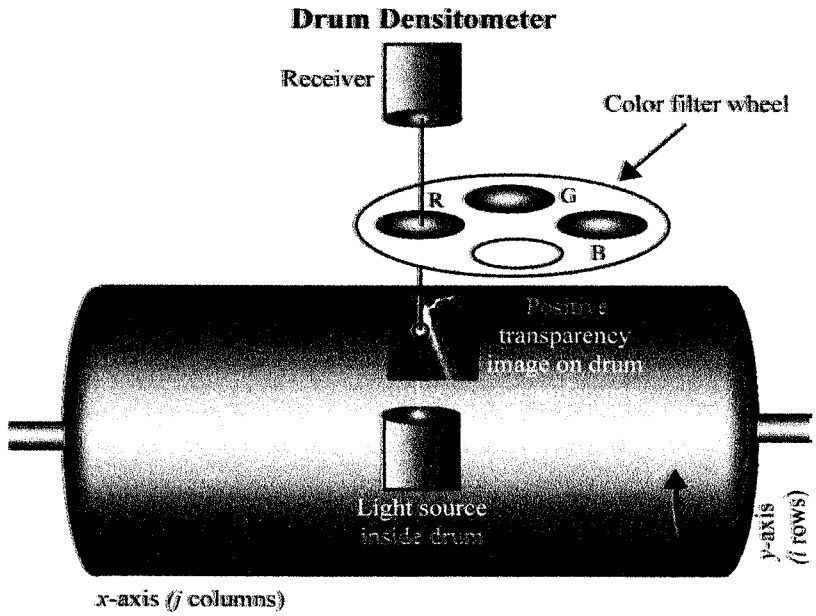


Figure 4-5: Schematic of drum densitometer (From Jensen, 2005).

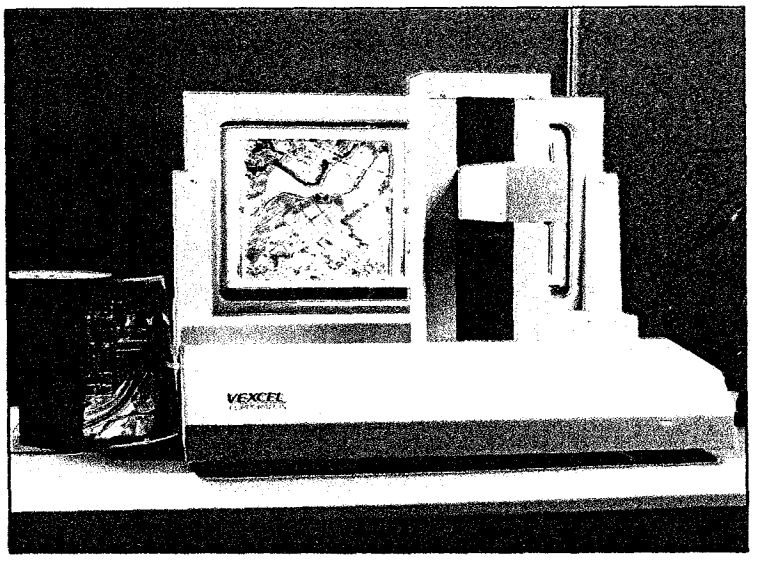


Figure 4-6: Video densitometer (From Jensen, 2005).

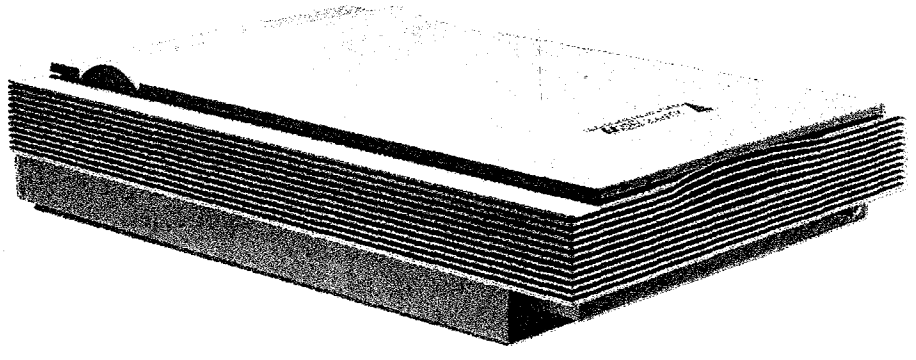


Figure 4-7: Video densitometer (From Jensen, 2005).

Theoretically, the densitometer can output the characteristics at each i,j location in the photograph in term of transmittance, opacity, or density. The ability of a portion of a developed films to pass light is called its transmittance ($\tau_{i,j}$).The transmittance at the location i,j in the photograph is :

$$\tau_{i,j} = \frac{\text{light pasing through the film}}{\text{total incident light}} \quad (4\text{-}2)$$

There is an inverse relationship between transmittance and the opacity of an area on the film. An area in the film that is very opaque does not transmit light well. Opacity ($O_{i,j}$) is the reciprocal of transmittance:

$$O_{i,j} = \frac{1}{\tau_{i,j}} \quad (4\text{-}3)$$

It is common to use density ($D_{i,j}$), which is the common logarithm of opacity for digitization measure of choice:

$$D_{i,j} = \log_{10} O_{i,j} = \log\left(\frac{1}{\tau_{i,j}}\right) \quad (4\text{-}4)$$

For example, if 10 percent of the light can be transmittance through a film at a certain *i,j* location, transmittance is 1/10, opacity is 1/0.10 or 10, and the density is the common logarithm of 10 or 1.0. As the light source is scanned across the image, the continuous output from the receiver is converted to a series of discrete numerical values on a pixel-by-pixel basis. This analog-to-digital conversion process results in a matrix of values that is usually recorded in 8-bit bytes. These data are then stored on disk or tape for future analysis (Jensen, 2005). Table 4-1 summarized the relationship between digitizer scanning spot size (IFOV) measured in dots per inches (dpi) or micrometer and the pixel ground resolution at various scales of aerial photography or image.

Table 4-1: Relationship between digitizer scanning spot size (IFOV) and the pixel ground resolution at various scales of aerial photography or image.

Digitizer Detector IFOV		**Pixel Ground Resolution at Various Scales of Photography (m)**					
Dots per inch	**Micrometers**	**1:40,000**	**1:20,000**	**1:9,600**	**1:4,800**	**1:2,400**	**1:1,200**
100	254.00	10.16	5.08	2.44	1.22	0.61	0.30
200	127.00	5.08	2.54	1.22	0.61	0.30	0.15
300	84.67	3.39	1.69	0.81	0.41	0.20	0.10
400	63.50	2.54	1.27	0.61	0.30	0.15	0.08
500	50.80	2.03	1.02	0.49	0.24	0.12	0.06
600	42.34	1.69	0.85	0.41	0.20	0.10	0.05
700	36.29	1.45	0.73	0.35	0.17	0.09	0.04
800	31.75	1.27	0.64	0.30	0.15	0.08	0.04
900	28.23	1.13	0.56	0.27	0.14	0.07	0.03
1000	25.40	1.02	0.51	0.24	0.12	0.06	0.03
1200	21.17	0.85	0.42	0.20	0.10	0.05	0.03
1500	16.94	0.67	0.34	0.16	0.08	0.04	0.02
2000	12.70	0.51	0.25	0.12	0.06	0.03	0.02
3000	8.47	0.33	0.17	0.08	0.04	0.02	0.01
4000	6.35	0.25	0.13	0.06	0.03	0.02	0.008

Using DPI PM = (S/DPI)/39.37 PF = (S/DPI)/12

Using micrometer PM = (S x μm)0.000001 PF = (S x μm)0.00000328

where: PM = pixel size in meters, PF = pixel size in feet, S =photo scale

(From Jensen, 2005)

4.2 Digital Image Resolution

In remote sensing, for type distinct types of resolution must be considered:

A. Spectral resolution - the specific wavelength intervals that a sensor can record (Figure 4-8).

B. Spatial resolution - the area on the ground represented by each pixel (Figure 4-9).

C. Radiometric resolution – the number of possible data file value in each band indicated by the number of bits into which the recorded energy is divided (Figure 4-10).

D. Temporal resolution – how often a sensor obtains imagery of a particular area (Figure 4-11). In addition, Figure 4-12 illustrated all four of resolution.

Jensen (2007) mentioned that there are often trade-offs associated with the various resolutions that must be made when collecting remote sensing data. (Figure 4-13) Generally, the higher the temporal resolution requirement (e.g., monitoring hurricane every half-hour), the lower the spatial resolution requirement (e.g., the NOAA GOES weather satellite records images with 4 X 4 to 8 X 8 km pixels). Conversely, the higher the spatial resolution requirement (e.g., monitoring urban land-use with 1 X 1 m data), the lower the temporal resolution requirement (e.g., every 1 to 10 years).

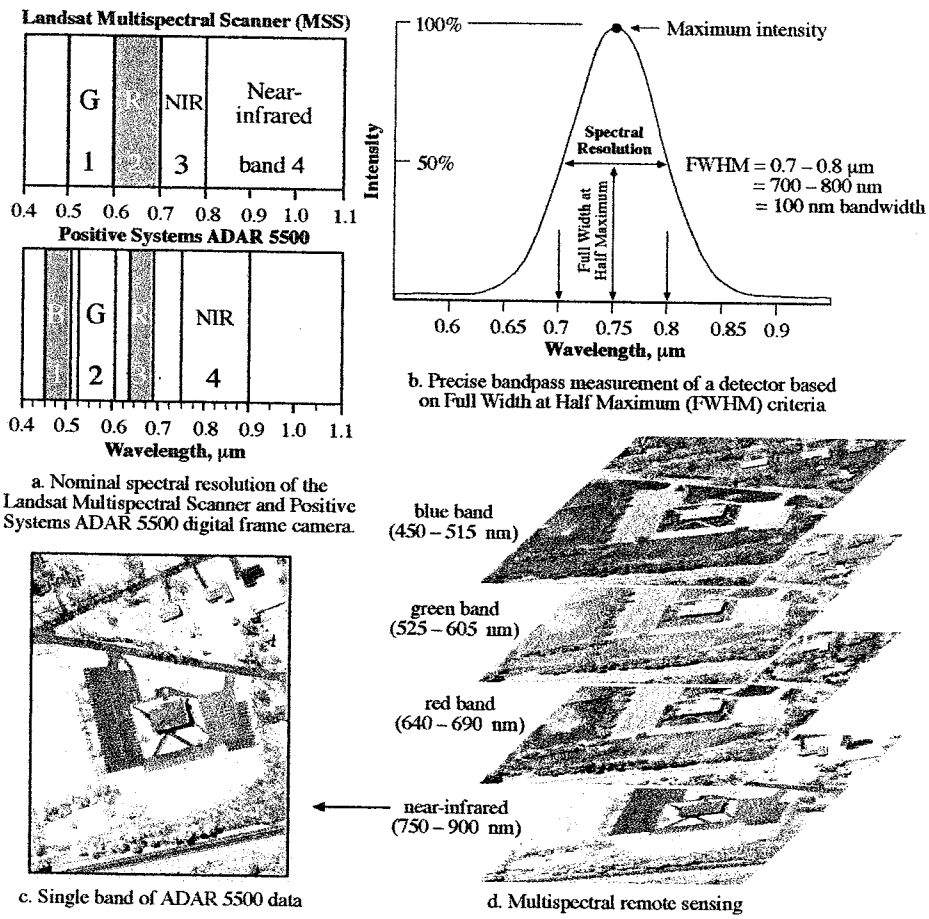


Figure 4-8: Characteristics of spectral resolution (From Jensen, 2005).

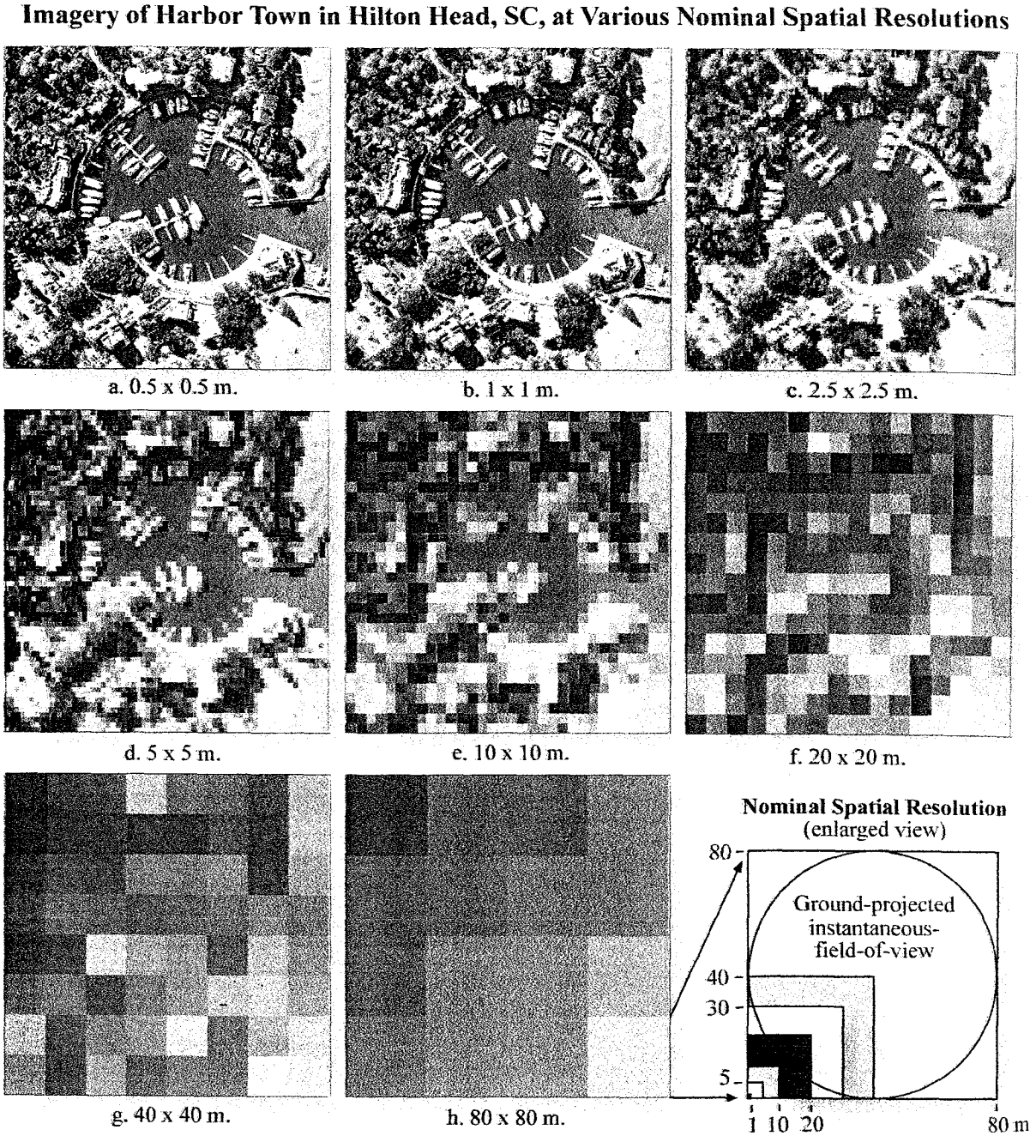


Figure 4-9: Characteristics of spatial resolution (From Jensen, 2005).

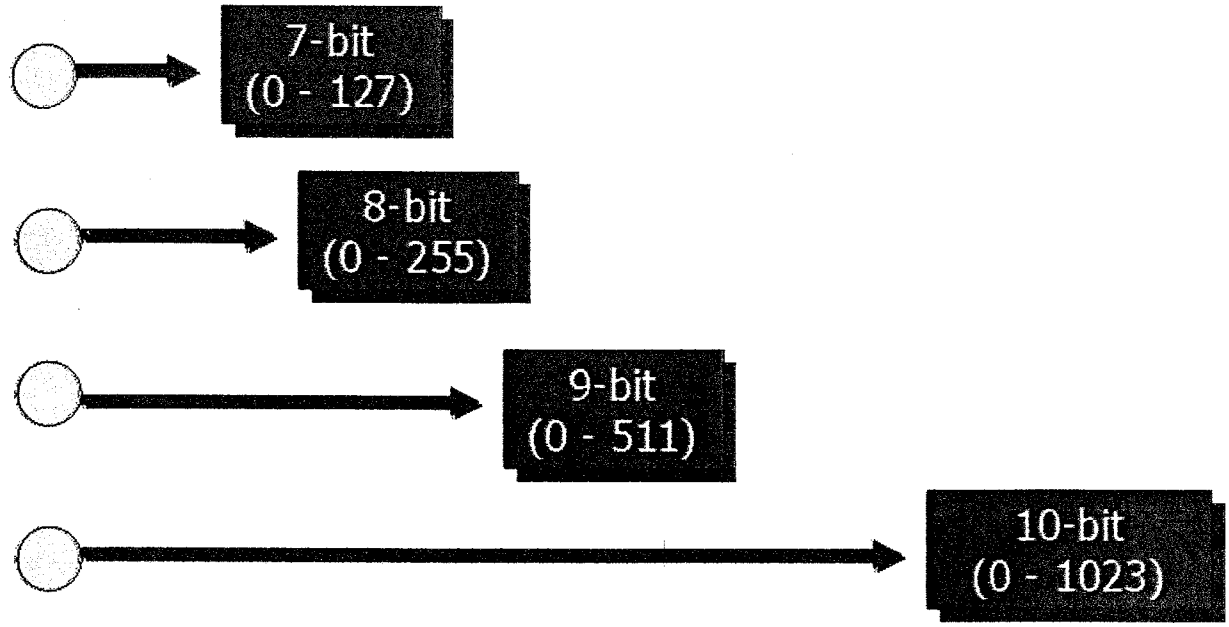


Figure 4-10: Characteristics of radiometric resolution (From Jensen, 2005).

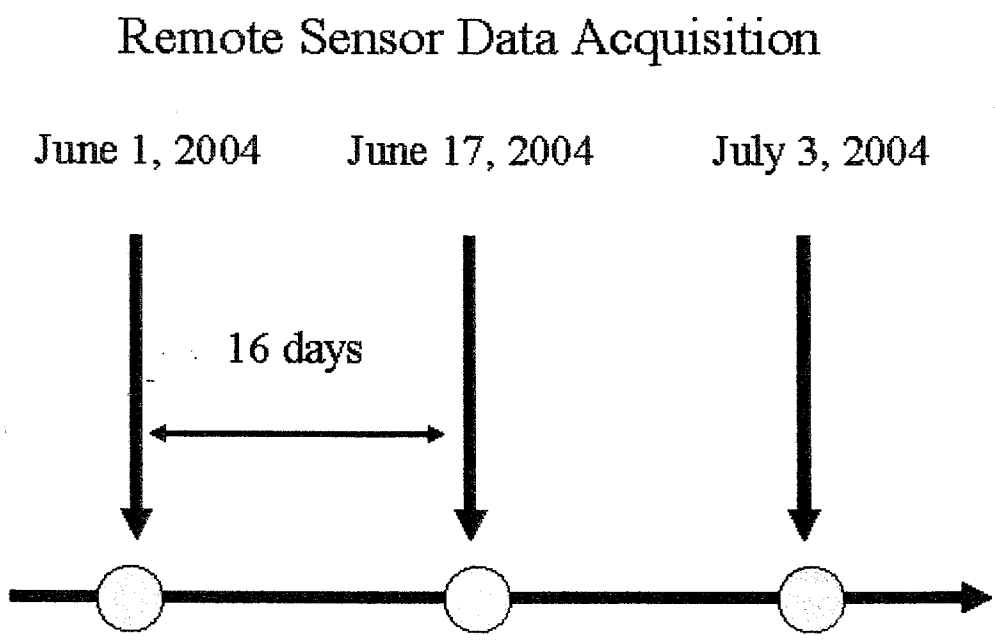


Figure 4-11: Characteristics of temporal resolution based on Landsat-5 (From Jensen, 2005).

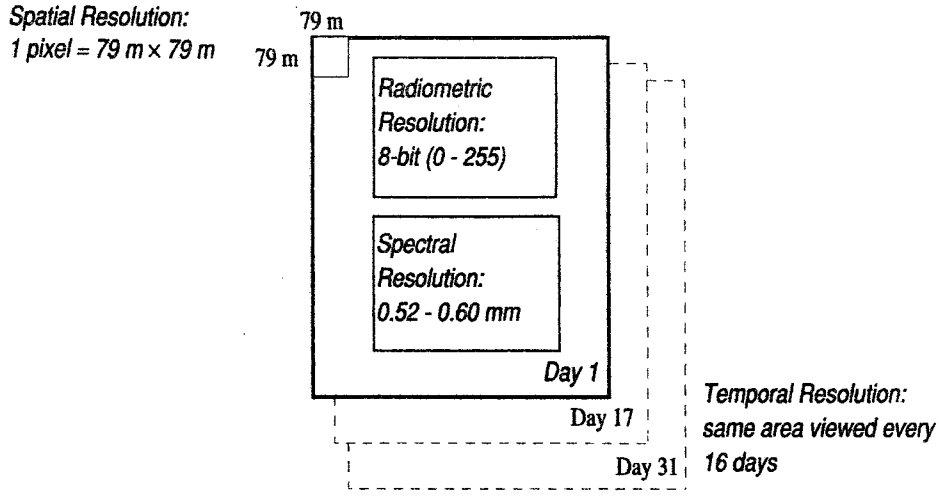


Figure 4-12: Four types of resolution of Landsat TM-Band 2 (From ERDAS, 2002.)

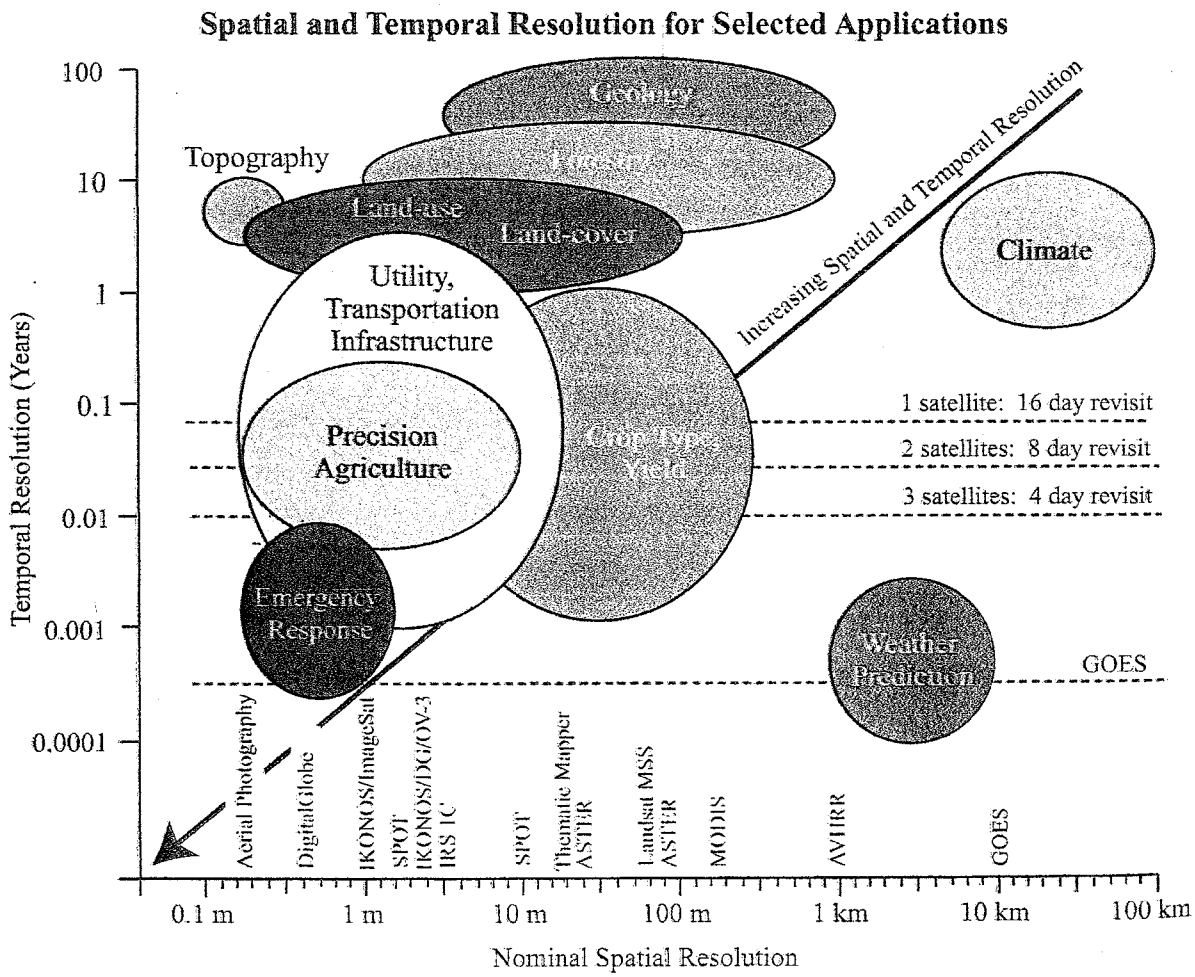


Figure 4-13: Spatial and temporal resolution considerations for selected application (From Jensen, 2007).

4.3 Digital Image Processing

The growth of all phase of digital computing technology has enabled the development of digital image processing technology.

Townshend and Justice (1981) defined image processing as a variety of transformation and manipulation for the ultimate purpose of improving our ability in extracting information from imagery. They recognized the following types of image processing as most relevant. Firstly, there is enhancement to aid information extraction by human interpreters. Secondly, they are various transformations to provide more suitable input for computer-based information extraction: these include derivation of textural measures (Haralick, 1979) and reductions in the dimensionality of data sets (Maxwell, 1976). Thirdly, there is information extraction in which computer-assisted methods are dominant. Finally, there are the many techniques involve in effective display of the find products. In this view, they also defined pre-processing as the creation of image products from the original sensed data with satisfactory geometric and radiometric properties for subsequent processing.

Bernstein (1983) stated that digital image processing involves the use of a computer to digitally manipulate or operate upon the matrix of numbers of a digital image for a particular objective.

Green (1983) stated that every digital image processing system consists of three majors elements: image acquisition, image processing, and image display.

Image Acquisition. Involves the conversion of a scene into a digital representation that can be processed by a digital computer. Image acquisition can be performed by a sensor system specially designed to view a scene and provide a digital representation of that scene; image acquisition can also involve the conversion of image data from an existing medium (television, film, etc.) into digital representation.

Image Processing. Provide digital processing of one or more images to produce a desired result. The processing can range from simple enhancement of an individual image for improved display of scene detail to more complex processing involving several component images (e.g., multispectral classification) or several hundred component images (e.g., a photomosaic of a large region of a planet).

Image Display. Provide for generation of an output that can be seen by a human observer. The display can be achieved by using volatile display techniques on a video monitor or in hard-copy from (e.g., film). This element provides the required conversion of digital data into some analog form (e.g., video signals or film products) for viewing

In practically, a typical digital image processing laboratory consists of computer systems and peripheral devices (Jensen, 2005) as shown in Figure 4-14.

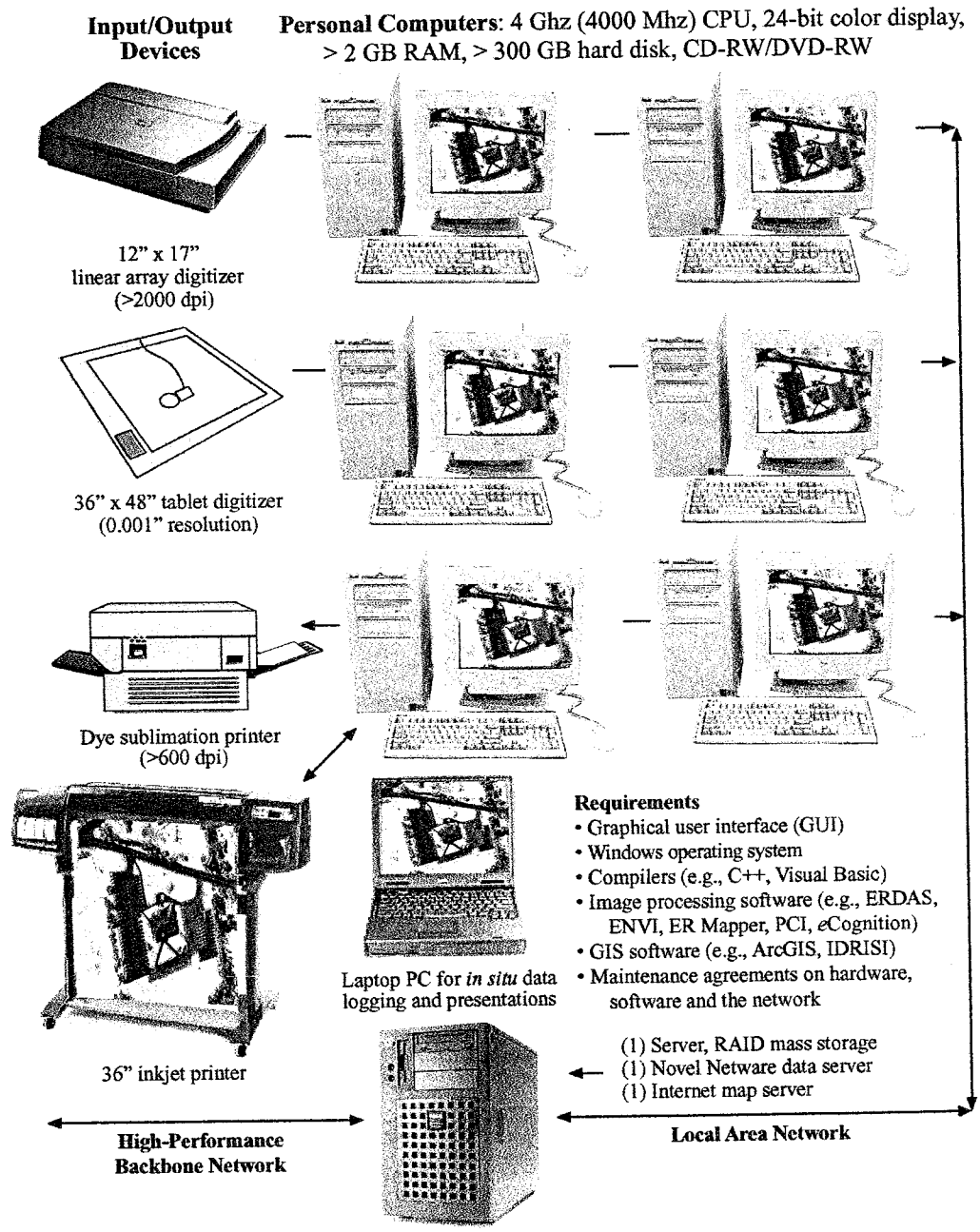


Figure 4-14: A typical digital image processing laboratory (From Jensen, 2005).

The most important functions performed using digital image processing systems based on Jensen (2005) was shown in Table 4-2 included:

- preprocessing,
- display and enhancement,
- information extraction,
- photogrammetric information extraction,
- metadata and image map lineage documentation,
- image and map cartographic composition,
- geographic information system,
- integrated image processing and GIS, and
- utilities.

In addition, digital image processing software can be categorized into two domains, namely, commercial firms and public agencies. Jensen (2005) summarized the major functions of selected commercial and digital image processing systems as shown in Table 4-3 and Table 4-4, respectively.

Table 4-2: Image processing functions in quality digital image processing systems

A. Preprocessing (Radiometric and Geometric)
1. Radiometric correction of error introduced by the sensor system electronics/optics and/or environmental effects.
2. Geometric correction (image-to-image registration or image-to-map rectification)
B. Display and Enhancement
3. B&W computer display (8 bits)
4. Color composite computer display (24 bits)
5. Magnification and reduction, roaming
6. Contrast manipulation (linear, nonlinear)
7. Color Space Transformation (RGB to HIS)
8. Image algebra (e.g., band ratio, image differencing, NDVI, SAVI, Kauth-Thomas, EVI)
9. Spatial filtering (e.g., low-pass, high-pass, bandpass)
10. Edge enhancement (e.g., Kirsh, Laplacian, Sobel)
11. Principal component
12. Texture transform (e.g., min-max, texture spectrum, fractal dimension, geostatistical)
13. Frequency transformations (e.g., Fourier, Walsh)
14. Digital Elevation Models (e.g., interpolation via IDW or Kriging, analytical hill shading, calculation of slope, aspect)
15. Three-dimensional transformations (e.g., image draping over DEM)
16. Image Animation (e.g., movies, change detection)
C. Information Extraction
17. Pixel brightness value
18. B&W or color slice
19. Transect (spatial and spectral)
20. Univariate and multivariate statistical analysis (e.g. mean, covariance, correlation, geostatistical)
21. Feature (band) selection
22. Supervised classification
23. Object-oriented image segmentation and classification
24. Incorporation of ancillary data during classification

Table 4-2: (continued)

C. Information Extraction
25. Expert system image analysis
26. Neural network image analysis
27. Fuzzy logic classification
28. Hyperspectral classification
29. Radar image processing
30. Accuracy assessment
31. Change detection
D. Photogrammetric Information Extraction
32. Soft-copy production of orthoimages
33. Soft-copy extraction of digital elevation models
34. Soft-copy extraction of planimetric detail
E. Metadata and Image/Map Lineage Documentation
Metadata
Complete image and GIS file processing history
F. Image and Map Cartographic Composition
Scaled Postscript level II output of images and maps
Geographic Information Systems (GIS)
Raster (image) based GIS
Vector (polygon) based GIS
Integrated Image Processing and GIS
Complete image processing systems (functions 1 through 37 plus utilities)
Complete image processing systems and GIS (functions 1 through 44 plus utilities)
Utilities
Network (e.g., local area network, Internet)
Image compression (single image, video)
Import and export of various file formats

From Jensen (2005)

Table 4-3: Major Functions of Selected Commercial Digital Image Processing Systems

Systems	Operating System	Prepro-cessing	Display & Enhancement	Information Extraction	Soft-copy Photogram-metry	Lineage	Image/Map Cartography	GIS	IP/ GIS
Commercial									
ACORN	Windows	●							
AGIS	Windows							●	○
Applied Analysis subpixel processing	Windows	●	○	●					○
ArcGIS Image and Feature Analyst	W/UNIX	●	●	●	●		●	●	●
ATCOR2	IDL	●		○					
AUTOCAD 2004	W/UNIX	●	●	○			○	●	
BAE Systems SOCET Set	W/UNIX	●	●	●	●			●	●
Blue Marble	W/UNIX	●	●				●		
EarthView	Windows	●	●	●					
*e*Cognition	Windows	●	●	●				○	○
EIDETIC Earthscope	Windows	○	○	○					

Table 4-3: (continued)

Systems	Operating System	Prepro-cessing	Display & Enhancement	Information Extraction	Soft-copy Photogram-metry	Lineage	Image/Map Cartography	GIS	IP/ GIS
ENVI	W/U/M/ IDL	●	●	●	○	○	●	○	○
DIMPLE	Mac	●	●	●			●	○	
Dragon	Windows	●	●	●					
ERDAS Imagine (Leica Geosystems)	W/UNIX	●	●	●	●	●	●	●	●
ER-Mapper	W/UNIX	●	●	●	●		●	●	●
FullPixelSearch	Mac	●	●	●					
GENASYS	W/UNIX	●	●	●			●	●	●
Global Lab Image	Windows		●	○					
GRASS	UNIX	●	●	●		●	●	●	●
IDRISI	Windows	●	●	●			●	●	●
ImagePro	Windows	●	●	●					
Intelligent Library Solution	UNIX	●	●			●	●		
Intergraph	W/UNIX	●	●	●	●	●	●	●	●

Table 4-3: (continued)

Systems	Operating System	Prepro-cessing	Display & Enhancement	Information Extraction	Soft-copy Photogram-metry	Lineage	Image/Map Cartography	GIS	IP/ GIS
MapInfo	W/UNIX		○			●	●	●	●
MrSID	W/U/M	●	●						
NOeSYS	W/Mac	○	●						
PCI Geomatica	W/UNIX/	●	●	●	●	●	●	●	●
Photoshop	W/U/M	○	●	○					
R-WEL	Windows	●	●	●	●		●	●	●
RemoteView	Windows	●	●						
MacSadie	Mac	●	●	●					
TNTmips	W/UNIX	●	●	●	●	●	●	●	●
OrthoView	UNIX	●	●						
VISILOG	W/UNIX	●	●	●					

Form Jensen (2005)

***Table 4-4:** Major Functions of Selected Public Digital Image Processing Systems*

Systems	Operating System	Prepro-cessing	Display & Enhancement	Information Extraction	Soft-copy Photogram-metry	Lineage	Image/Map Cartography	GIS	IP/ GIS
C-Coast	Windows		●	●					
Cosmic VICAR-IBIS	UNIX	●	●	●			●	●	●
NOAA	UNIX	○	○						
MultiSpec	Mac/W	●	●	●					
NASA ELAS (DIPIX, Dtastar)	UNIX	●	●	●		●	●	●	●

Form Jensen (2005)

4.4 An Ideal Step of Digital Image Processing

Campbell (1987) suggested an idealized sequence for a digital analysis as shown in Figure 4-15. The details of each step were summarized in the following sections:

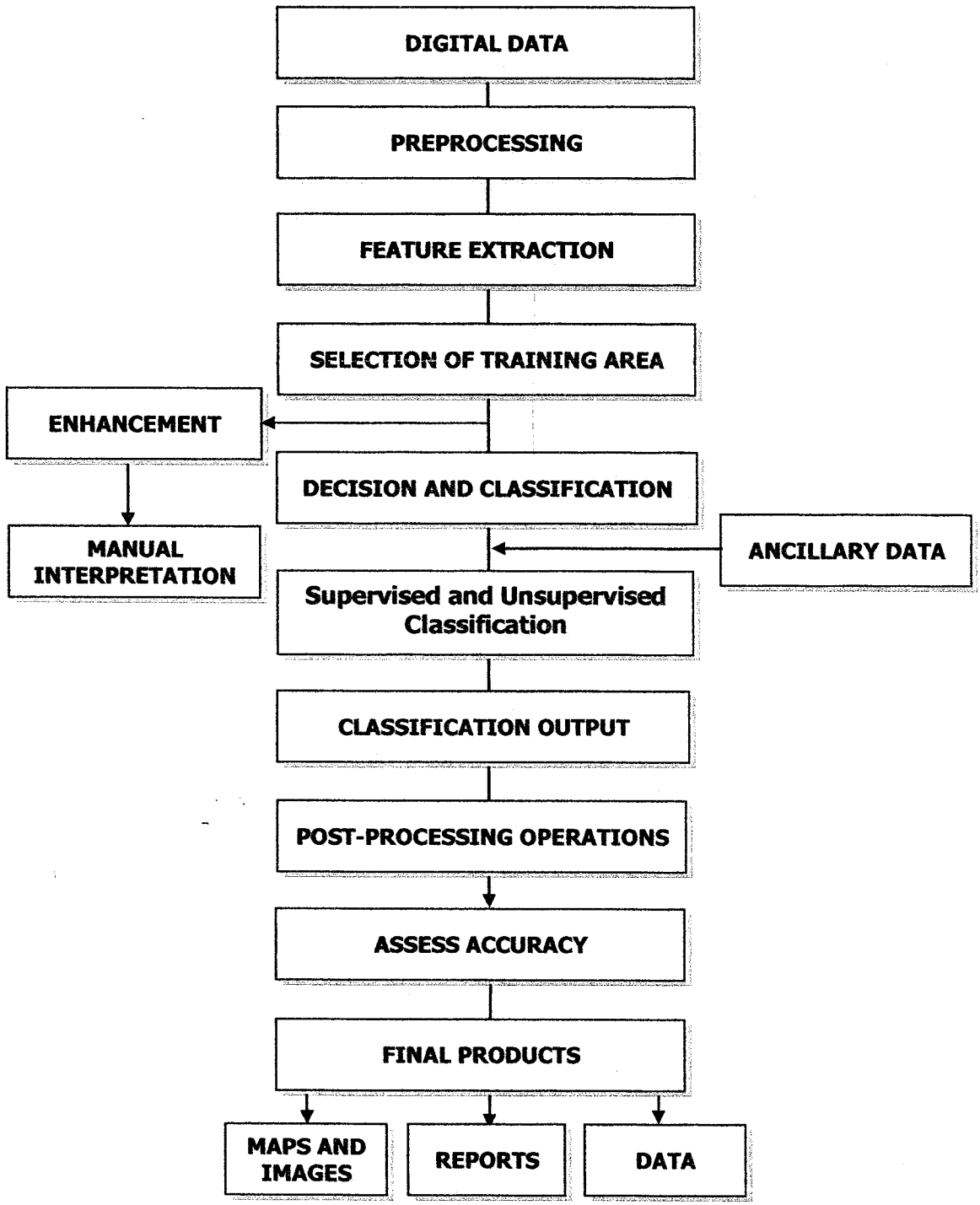


Figure 4-15: Idealized sequence for digital image analysis (From Campbell, 1987).

A. Preprocessing

Preprocessing consists of those operations that prepare data for subsequent analysis, usually by attempts to correct or compensate for systematic errors. Three classes of preprocessing operations can be defined.

The first are those that simply display, or summarize, the data as a means of inspecting characteristics and quality. These operations present histograms, scatterograms, or statistical summaries that permit the operator to assess image quality and thereby determine subsequent preprocessing steps (if any) that may be necessary.

A second group of preprocessing operations are those that compensate for *radiometric errors*. These errors results from defects in sensor operations atmospheric absorption and scattering, variations in scan image, variations in illumination of the scene, and system noise.

A third class of preprocessing operations addresses errors in image geometric. *Geometric errors* are positional errors in relationships between image representations of ground features and their actual geographic relationships as they occur on the earth's surface. As a result, geometric preprocessing shifts positions of pixels to approximate their true ground positions.

B. Feature Extraction

After preprocessing is complete, the analyst may use feature extraction to reduce the dimensionality of the data. Feature extraction is the process of isolating the most useful components of the data for further study while discarding the less useful aspects (errors, noise, etc.). Feature extraction reduces the number of variables that must be examined, thereby saving time and resource.

C. Image Enhancement

As this point the analyst may choose to apply techniques to improve the visual appearance of the image. Image enhancement prepares the image for subsequent manual interpretation. Because image enhancement often drastically alters the original numeric data, it is normally used only for visual interpretation and is not used if further numeric analysis is to be undertaken.

E. Decision and Classification

There are several alternative approaches to decision and classification. In this step, data that have subjected to processing and feature extraction are subjected to quantitative analysis to assign individual pixels to specific classes. That is, individual pixels are groups may correspond to regions on the ground that have common geologic or biologic properties, and therefore may be useful to the analyst as a means of deriving information concerning such subjects as land use, water quality, geology, or plant vigor.

In general, there are two approaches to image classification: supervised classification and unsupervised classification.

Supervised Classification uses information derived from a few areas of known identity to classify the remainder of the image, consisting of those pixels of unknown identity. The analyst must carefully locate both on the ground and on the image areas that represent the class of interest to be represented on the final product. Once these areas, training areas or training fields, have been accurately located on the image, the classification procedure uses characteristics of these areas to assign pixels of unknown identity to each of the categories specified by the operator.

The alternate, ***unsupervised classification*** is the search natural groupings of pixels based upon their brightness in several spectral channels. The analyst then attempts to assign these natural classes to the users-defined categories (such as forest or water) that are of value the scientists and planner who will use the results.

Choice between supervised and unsupervised method must often be made on the basis of information at hand, or for the purposes of a specific study. Sometimes it is possible to combine characteristics of both methods , using the control provided by the supervised approach with the objectivity of the unsupervised strategy and its ability to identify the natural structure of spectral classes in an image. Use of both supervised and unsupervised method in a single procedure is known **as *hybrid classification.***

F. Access Accuracy

After classification is complete, it is necessary to evaluate its accuracy by comparing the categories on the classified images with areas of known identity on the ground. If such comparisons can be conducted for each category, the analyst can develop information that permits assessment of the categories that are most accurately classified and can document the quality of the results.

G. Final Product

The final results of the analyst consist of the maps (or images), data, and a report. The maps and images provide the user with knowledge of spatial distributions of the categories of interest and their patterns on the landscapes. Data are provided by tabulations of classes, and their areas. Finally, the report provides a clear account of the images examined, the procedures that were used, and the results of the accuracy assessment. These three components of the results provide the user with full information concerning the source data, method of analysis, the outcome, and its reliability.

Jensen (2005) stated that the analysis of remotely sensed data is performed using a variety of image processing techniques (Figure 4-16), including:

- analog (visual) image processing, and
- digital image processing.

Analog and digital analyses of remotely sensed data seek to detect and identify important phenomena in the scene. Once identified, the phenomena are usually measured, and the information is used in solving problems. Thus, both manual and digital analyses have the same general. However, the attainment of these goals may follow significantly different paths. Optimum results are often achieved using a synergistic combination of both visual and digital image processing.

Fundamental Image Analysis Tasks
- Detect, Identify, Measure
- Solve Problems

Application of the ***Multi*** concept
- Multispectral - Multifrequency - Multipolarization
- Multitemporal - Multiscale - Multidisciplinary

Use of ***Collateral Information***
- Literature - Laboratory spectra - Dichotomous keys - Prior probabilities
- Field training sites - Field test sites - Soil maps - Surficial geology maps

Analog (Visual) Image Processing	**Digital Image Processing**
Elements of Image Interpretation	***How the Elements of Image Interpretation Are Extracted or Used in Digital Image Processing***
• Grayscale tone (black to white)	• 8- to 12-bit brightness values, or more appropriately scaled surface reflectance or emittance
• Color (red, green, blue = RGB)	• 24-bit color look-up table display - Multiband RGB color composites - Transforms (e.g., intensity, hue, saturation)
• Height (elevation) and depth	• Soft-copy photogrammetry, LIDAR, radargrammetry, RADAR interferometry, SONAR
• Size (length, area, perimeter, volume)	• Soft-copy photogrammetry, radargrammetry, RADAR interferometry, measurement from rectified images
• Shape	• Soft-copy photogrammetry, radargrammetry, RADAR interferometry, landscape ecology spatial statistics (metrics), object-oriented image segmentation
• Texture	• Texture transforms, geostatistical analysis (e.g., kriging), landscape ecology metrics, fractal analysis
• Pattern	• Autocorrelation, geostatistical analysis, landscape ecology metrics, fractal analysis
• Shadow	• Soft-copy photogrammetry, radargrammetry, measurement from rectified images
• Site, using convergence of evidence	• Contextual, expert system, neural network analysis
• Association, using convergence of evidence	• Contextual, expert system, neural network analysis
• Arrangement, using convergence of evidence	• Contextual, expert system, neural network analysis

Figure 4-16: Analog (visual) and digital image processing image processing of remotely sensed data use the fundamental elements of image processing (From Jensen, 2005).

Chapter 5: Preprocessing in Digital Image Processing

Preprocessing consists of those operations that prepare data for subsequent analysis, usually by attempts to correct or compensate for systematic errors (Campbell ,1987). Three main preprocessing operations, that prepare data for subsequent analysis, include (1) image quality assessment and statistical evaluation, (2) radiometric correction and (3) geometric correction.

- **Image quality assessment and statistical evaluation.** Many remote sensing datasets contain high-quality, accurate data. Unfortunately, sometimes error (or noise) is introduced into the remote sensor data by: (a) the environment (e.g., atmospheric scattering), (b) random or systematic malfunction of the remote sensing system (e.g., an uncalibrated detector creates striping), or (c) improper airborne or ground processing of the remote sensor data prior to actual data analysis (e.g., inaccurate analog-to-digital conversion). Therefore, users for analyzing the digital remote sensor should first assess its quality and statistics characteristics. (Jensen, 2005).

Radiometric and geometric correction. Remote sensing systems do not function perfectly. Also, the Earth's atmosphere, land, and water are complex and do not lend themselves well to being recorded by remote sensing devices that have constraints such as spatial, spectral, temporal, and radiometric resolution. Consequently, error creeps into the data acquisition process and can degrade the quality of the remote sensor data collected. The two most common types of error encountered in remotely sensed data are radiometric and geometric. Radiometric correction attempts to improve the accuracy of spectral reflectance, emittance, or back-scattered measurements obtained using a remote sensing system. While, geometric correction is concerned with placing the reflected, emitted, or back-scattered measurements or derivative products in their proper planimetric (map) location so they can be associated with other spatial information in a geographic information system (GIS) or spatial decision support system (SDSS).

Radiometric and geometric corrections of remotely sensed data are normally referred to as preprocessing operations because they are performed prior to information extraction. Image preprocessing hopefully produces a corrected image that is as close as possible, both radiometrically and geometrically, to the true radiant energy and spatial characteristics of the study area at the time of data collection. Internal and external errors must be identified to correct the remotely sensed data (Jensen, 2005):

- Internal errors are introduced by the remote sensing system. They are generally systematic (predictable) and may be identified and then corrected based on prelaunch or in-flight calibration measurements.
- External errors are introduced by phenomena that vary in nature through space and time. External variables that can cause remote sensor data to exhibit radiometric and geometric error include the atmosphere, terrain elevation, slope, and aspect. Some external errors may be corrected by relating empirical ground observations (i.e., radiometric and geometric ground control points) to sensor measurements.

In this chapter these preprocessing operations will be here explained and demonstrated more in detail.

5.1 Image Quality Assessment and Statistical Evaluation

Basically, Jensen (2005) suggested that image quality assessment and statistical evaluation is normally accomplished by:

- looking at the frequency of occurrence of individual brightness values in the image displayed in a histogram formats,
- viewing on a computer monitor individual pixel brightness values at specific locations or within a geographic area,
- computing fundamental univariate descriptive statistics to determine if there are unusual anomalies in the image data,
- computing multivariate statistics to determine the amount of between-band correlation (e.g., to identify redundancy).

5.1.1 Histogram Characteristics of Remote Sensor Data

The *histogram* is a useful graphic representation of the information content of a remotely sensed image. Histograms for each band of imagery are often displayed and analyzed in many remote sensing investigations because they provide the analyst with an appreciation of the quality of the original data (e.g., whether it is low in contrast, high in contrast, or multimodal in nature). In fact, many analysts routinely provide before and after histogram of the imagery to document the effect of applying an image enhancement. It is instructive to review how a histogram of a single band of imagery, k, composed of i rows and j columns with a brightness value BV_{ijk} at each pixel location is constructed. (Jensen 2005). Typical histogram of symmetric and skewed distribution is shown in Figure 5-1

The range of quantized of a band of image is plotted on the x axis, while the frequency of occurrence of each values is displayed on the y axis. In fact, the peaks in the histogram correspond

to the dominant types of land use. Also, the range of brightness in each band suggest about the contrast of the data. The example of histogram was shown Figure 5-2.

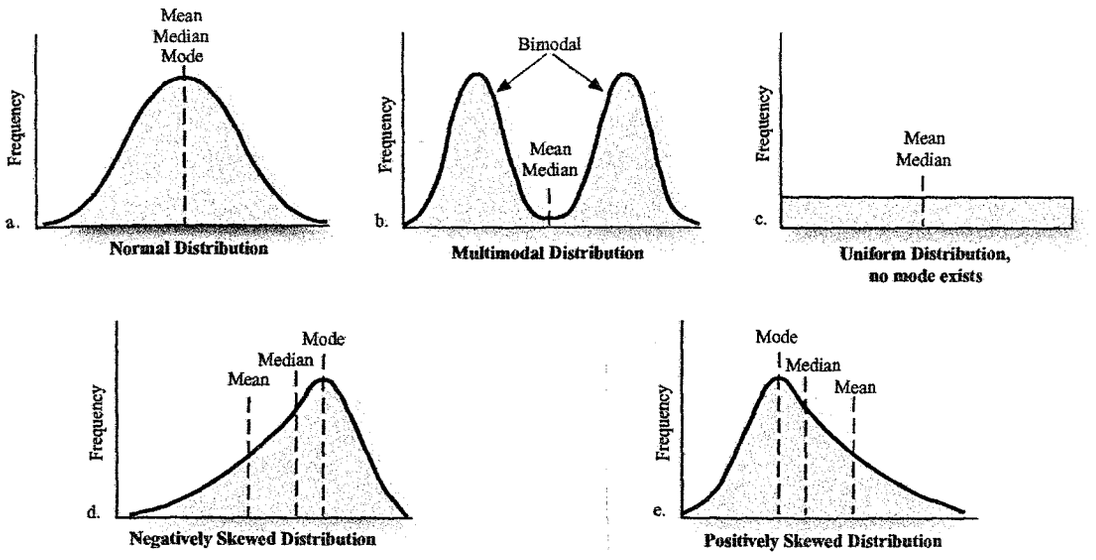


Figure 5-1: Histograms of symmetric and skewed distribution. (From Jensen, 2005)

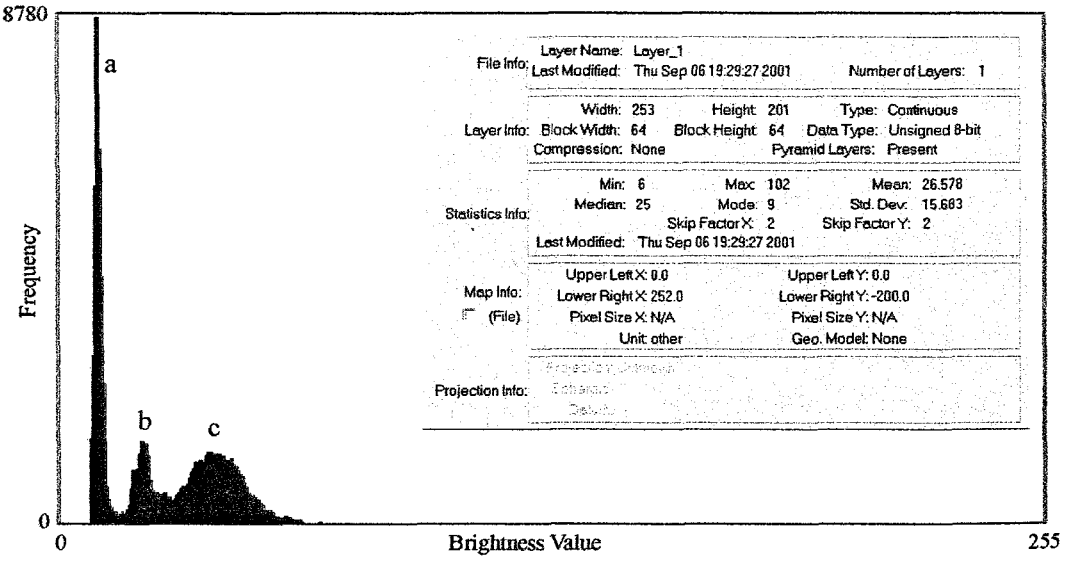
5.1.2 Image Metadata

Metadata is **"data or information about data"**. Most quality digital image processing systems read, collect, and store metadata about a particular image or sub-image. It is important that the image analyst have access to this metadata information. In the most fundamental instance, metadata might include:

- file name
- date of last modification
- level of quantization
- number of rows and columns
- number of bands
- univariate statistics
- multivariate statistic
- geo-referencing performed (if any)
- pixel dimension.

The metadata information helps the analyst understand the content of remote sensor data.

a. Landsat Thematic Mapper Band 4 Image of Charleston, South Carolina (contrast stretched).



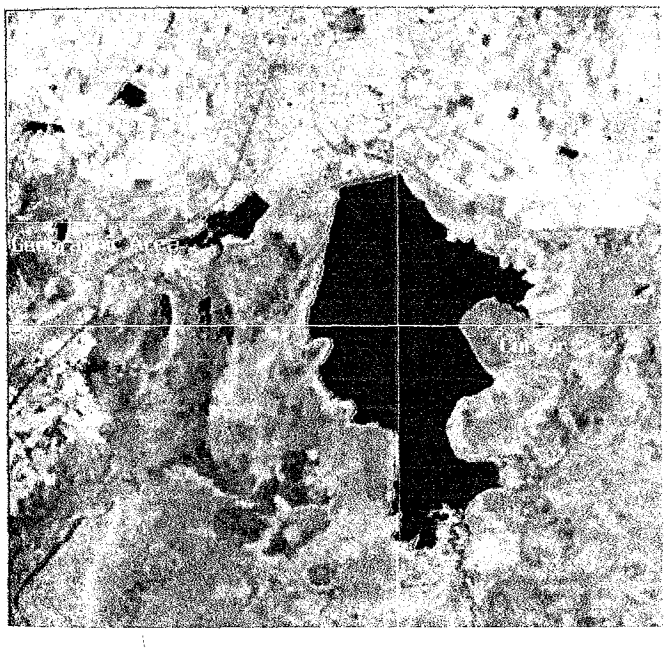
a. Histogram of the original Landsat Thematic Mapper Band 4 imagery.

Figure 5-2: Histogram of a single band of Landsat Thematic Mapper data of Charleston, SC, obtaining on November 9, 1982. Peaks in the histogram correspond to dominant types of land cover in the image, including: (a) open water pixels, (b) coastal wetland, and (c) upland (From Jensen, 2005).

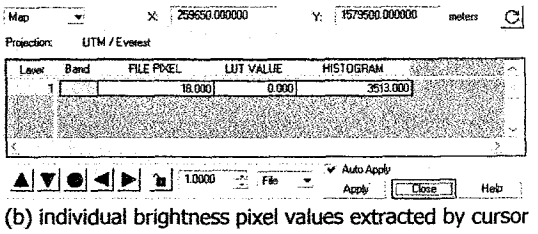
5.1.3 Viewing Individual Pixel Brightness Value

Viewing individual pixel brightness values in a remotely sensed image is one of the most useful methods for assessing the quality and information content of the data. Virtually all digital image processing systems allow the analyst to:

- use a mouse-controlled *cursor* (cross-hair) to identify a geographic location in the image (at a particular row and column or geographic x, y coordinate) and display its brightness value in *n* bands
- display the individual brightness values of an individual band in a *matrix* (raster) format. Example of viewing individual pixel brightness values shown in Figure 5-3



a. Landsat-TM Band 4 in 2004: Non Din Deang, Buriram, Thailand



(b) individual brightness pixel values extracted by cursor

Row	0	1	2	3	4	5	6	7	8	9	10	11
0	90	87	78	70	67	64	62	67	77	84	86	88
1	84	77	71	67	63	62	63	72	80	86	88	89
2	75	69	65	62	61	62	63	72	80	84	86	89
3	66	62	59	60	63	65	64	69	74	82	89	92
4	67	61	58	60	64	67	66	67	72	82	91	95
5	66	61	59	61	64	64	64	66	73	85	95	98
6	61	62	62	63	66	66	66	67	76	88	95	97
7	61	62	61	65	68	69	69	71	80	91	96	97
8	69	66	63	68	70	69	68	71	81	92	96	97
9	78	74	70	71	71	70	68	73	83	92	96	98
10	82	81	77	72	69	69	70	76	84	91	95	99
11	81	81	78	73	66	67	75	80	84	89	93	97
12	78	80	76	69	63	65	76	83	86	87	87	88
13	79	78	73	66	62	65	76	84	88	88	88	88
14	79	76	71	64	61	68	82	88	87	87	87	87
15	79	77	70	62	61	71	84	88	86	86	86	85
16	79	77	70	62	63	74	85	87	84	85	86	83
17	77	74	71	68	67	77	84	84	82	84	86	82
18	76	73	74	79	76	80	82	82	82	83	86	81
19	74	72	77	83	81	82	81	85	89	90	85	81
20	73	73	79	84	81	81	79	85	91	91	86	81
21	72	74	82	85	82	77	78	87	91	90	86	82

c. Brightness values in a geographic area for individual band

Figure 5-3: Cursor evaluation of individual pixel brightness values.

5.1.4 Univariate Descriptive Image Statistics

Most digital image processing system can carry out univariate and multivariate analysis of single and multiple band remotely sensed data. The major analyses are measure of central tendency and dispersion of remote sensor data.

A. Measures of Central Tendency in Remote Sensor Data

Mode: The mode is the value that occurs most frequently in a distribution and is usually the highest point on the curve (histogram). It is common, however, to encounter more than one mode in remote sensing data set (See Figure 5-2)

Median: The median is the value midway in the frequency distribution. One-half of the area below the distribution curve is to the right of the median, and one-half is to the left.

Mean: The mean is the arithmetic average and is defined as the sum of all brightness value observations divided by the number of observations. It is the most commonly used measure of central tendency. The mean (μ_k) of a single band of imagery composed of n brightness values (BV_{ik}) is computed using the formula:

$$\mu_k = \frac{\sum_{i=1}^{n} BV_{ik}}{n} \quad (5\text{-}1)$$

The sample mean, μ_k, is an unbiased estimate of the population mean. For symmetrical distributions, the sample mean tends to be closer to the population mean than any other unbiased estimate (such as the median or mode). Unfortunately, the sample mean is a poor measure of central tendency when the set of observations is skewed or contains an outlier. If a peak (mode) falls to the right of the mean, the frequency distribution is negatively skewed. If a peak (mode) falls to the left of the mean, the frequency distribution is positively skewed (Jensen, 2005). (See Figure 5-1 d and e)

B. Measures of Dispersion: Range

Measures of the dispersion about the mean of a distribution provide valuable information about the image.

Range: The range of a band of imagery ($range_k$) is computed as the difference between the maximum (max_k) and minimum (min_k) values; that is, $range_k = max_k - min_k$. Unfortunately, when the minimum or maximum values are extreme or unusual observation (i.e., possibly *data blunders*), the range can be a misleading measure of dispersion. When unusual values are not encountered, the range is a very important statistic often used in image enhancement functions such as min–max contrast stretching.

Variance: The variance of a sample is the average squared deviation of all possible observations from the sample mean. The variance of a band of imagery, var_k, is computed:

$$\text{var}_k = \frac{\sum_{i=1}^{n} \left(BV_{ik} - \mu_k\right)^2}{n} \quad (5\text{-}2)$$

The numerator of the expression is *the corrected sum of squares* (SS). If the sample mean (μk) were actually the population mean, this would be an accurate measurement of the variance. Unfortunately, there is some underestimation because the sample mean was calculated in a manner that minimized the squared deviations about it. Therefore, the denominator of the variance equation is reduced to $n-1$, producing a larger, unbiased estimate of the sample variance:

$$\text{var}_k = \frac{SS}{n-1} \quad (5\text{-}3)$$

Standard deviation: The *standard deviation* is the positive square root of the variance. The standard deviation of the pixel brightness values in a band of imagery, s_k, is computed as:

$$s_k = \sigma_k = \sqrt{\text{var}_k} \quad (5\text{-}4)$$

A small standard deviation suggests that observations are clustered rightly around a central value. Conversely, a large standard deviation indicates that values are scattered widely about mean. The total area underneath a distribution curve is equal to 1.00 or 100 %. The area under the normal curve for various standard deviations was shown in Figure 5-4.

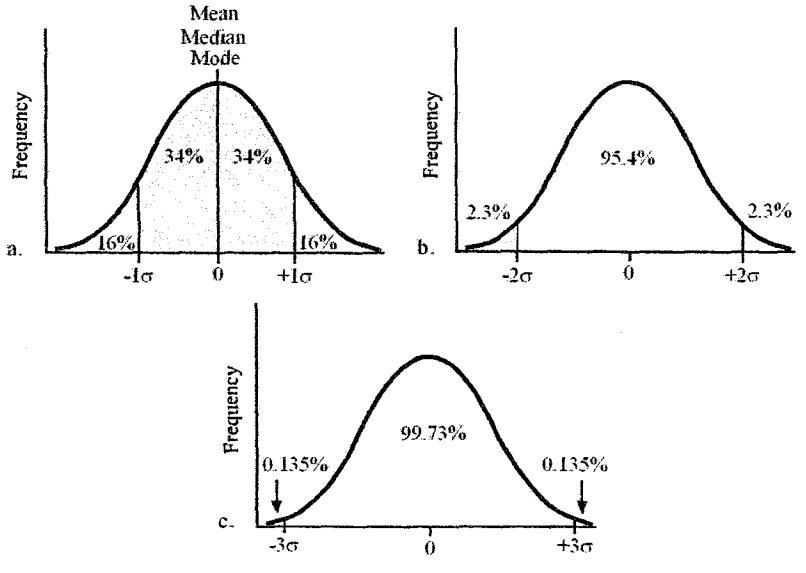


Figure 5-4: The areas under the normal curve for various standard deviations. (From Jensen, 2005)

To interpret variance and standard deviation, analysts should not attach a significant to each numerical value, but should compare one variance or standard deviation to another. The sample having the largest variance or standard deviation has the greater spread among the brightness value of the observations.

The example of simple univariate statistics of a hypothetical dataset of remote sensor data set are shown in Table 5-1.

Table 5-1: Hypothetical Dataset of Brightness Values

Pixel	Band 1	Band 2	Band 3	Band 4
(1,1)	130	57	180	205
(1,2)	165	35	215	255
(1,3)	100	25	135	195
(1,4)	135	50	200	220
(1,5)	145	65	205	235
Mean (μ_k)	135	46.40	187	222
Variance (var_k)	562.50	264.80	1007	570
Standard deviation (s_k)	23.71	16.27	31.4	23.87
Minimum (min_k)	100	25	135	195
Maximum (max_k)	165	65	215	255
Range (BV_r)	65	40	80	60

C. Measures of Distribution (Histogram) Asymmetry and Peak Sharpness

Skewness: Skewness is a measure of the asymmetry of a histogram and is computed using the formula:

$$skewness_k = \frac{\sum_{i=1}^{n}\left(\frac{BV_{ik} - \mu_k}{s_k}\right)^3}{n} \qquad (5\text{-}5)$$

A perfectly symmetric histogram has a *skewness* value of zero.

Kurtosis: A perfectly normal distribution (histogram) has zero *kurtosis*. The greater the positive kurtosis value, the sharper the peak in the distribution when compared with a normal histogram. Conversely, a negative kurtosis value suggests that the peak in the histogram is less sharp than that of a normal distribution. *Kurtosis* is computed using the formula:

$$kurtosis_k = \left[\frac{1}{n} \sum_{i=1}^{n} \left(\frac{BV_{ik} - \mu_k}{s_k} \right)^4 \right] - 3 \qquad (5\text{-}6)$$

5.1.5 Multivariate Descriptive Image Statistics

Remote sensing research is often concerned with the measurement of how much radiant flux is reflected or emitted from an object in more than one band (e.g., in red and near-infrared bands). It is useful to compute *multivariate* statistical measures such as *covariance* and *correlation* among the several bands to determine how the measurements covary. Also, variance–covariance and correlation matrices are used in remote sensing principal components analysis (PCA), feature selection, classification and accuracy assessment.

A. Covariance in Multiple Bands of Remote Sensor Data

The different remote sensing-derived spectral measurements for each pixel often change together in some predictable fashion. If there is no relationship between the brightness value in one band and that of another for a given pixel, the values are mutually independent; that is, an increase or decrease in one band's brightness value is not accompanied by a predictable change in another band's brightness value. Because spectral measurements of individual pixels may not be independent, some measure of their mutual interaction is needed. This measure, called the *covariance*, is the joint variation of two variables about their common mean (Jensen, 2005). To calculate covariance: we first compute the *corrected sum of products* (SP) defined by the equation:

$$SP_{kl} = \sum_{i=1}^{n} \left(BV_{ik} - \mu_k \right) \left(BV_{il} - \mu_l \right) \qquad (5\text{-}7)$$

Where:

BV_{ik} is ith brightness value in band k

BV_{il} is ith brightness value in band l

n is total number of pixels in an array

μ_k is mean of band k

μ_l is mean of band l.

It is more efficient to use the following formula to arrive at the same result:

$$SP_{kl} = \sum_{i=1}^{n} \left(BV_{ik} \times BV_{il}\right) - \frac{\sum_{i=1}^{n} BV_{ik} \sum_{i=1}^{n} BV_{il}}{n} \quad (5\text{-}8)$$

This quantity is called the *uncorrected sum of products*. A simple variance was calculated by dividing the corrected sums of squares (SS) by *(n – 1)*, *covariance* is calculated by dividing SP by *(n – 1)*. Therefore, the covariance between brightness values in bands *k* and *l*, cov_{kl} is equal to:

$$\text{cov}_{kl} = \frac{SP_{kl}}{n-1} \quad (5\text{-}9)$$

Table 5-2: shown covariance of a hypothetical dataset from Table 5-1 in variance-covariance matrix format.

Table 5-2: Variance-covariance matrix of the sample data.

	Band 1	**Band 2**	**Band 3**	**Band 4**
Band 1	562.25			
Band 2	135	264.80		
Band 3	718.75	275.25	1007.50	
Band 4	537.50	64	663.75	570

B. Correlation between Multiple Bands of Remotely Sensed Data

To estimate the degree of interrelation between variables in a manner not influenced by measurement units Pearson's product-moment correlation coefficient (r) is commonly used (Jensen, 2005). The correlation between two bands of remotely sensed data, r_{kl}, is the ratio of their covariance (cov_{kl}) to the product of their standard deviations ($s_k s_l$); thus:

$$r_{kl} = \frac{\text{cov}_{kl}}{s_k s_l} \quad (5\text{-}10)$$

Because the correlation coefficient is a ratio, correlation is a unitless number and ranges from +1 to -1. A correlation coefficient of +1 indicates a positive, perfect relationship between the brightness values in two of the bands (i.e., as one band's pixel increase in value, the other band's values also increase in a systematic fashion). Conversely, a correlation coefficient of -1 indicates

that two bands are inverse related (i.e., as brightness vales in one band increase, corresponding pixels in the other band systematically decreased in values) (Jensen, 2005). (See Table 5-3)

If we square the correlation coefficient (r_{kl}), we obtain the *sample coefficient of determination* (r^2), which expresses the proportion of the total variation in the values of "band *l*" that can be accounted for or explained by a linear relationship with the values of the random variable "band *k*." Thus a correlation coefficient (r_{kl}) of 0.70 results in an r^2 value of 0.49, meaning that 49% of the total variation of the values of "band *l*" in the sample is accounted for by a linear relationship with values of "band *k*" (Jensen, 2005).

Table 5-3: Correlation Matrix of the Sample Data

	Band 1	**Band 2**	**Band 3**	**Band 4**
Band 1	-			
Band 2	0.35	-	-	
Band 3	0.95	0.53	-	
Band 4	0.94	0.16	0.87	-

Based on Table 5-3, brightness values of band 1 are highly correlated with those of band 3 and 4, that is r >= 0.94. A high correlation suggests substantial redundancy in the information content among these bands could be deleted from the analysis to reduce subsequent computation. Conversely, the relatively lower correlation between band 2 and all other bands suggests that this band provides some type of unique information not found in other hands.

5.1.6 Feature Space Plots

Individual bands of remotely sensed data are often referred to as features in the pattern recognition literature. To truly appreciated how two bands (features) in a remote sensing dataset covary and if they are correlated or not, it is often useful to produce a two-band feature space plot.

A *two-dimensional feature space plot* extracts the brightness value for every pixel in the scene in two bands and plots the frequency of occurrence in a 255 by 255 feature space (assuming 8-bit data). The greater the frequency of occurrence of unique pairs of values, the brighter the feature space pixel (Jensen, 2005). An example of two dimensional feature space of the two bands of Landsat-TM was shown In Figure 5-5.

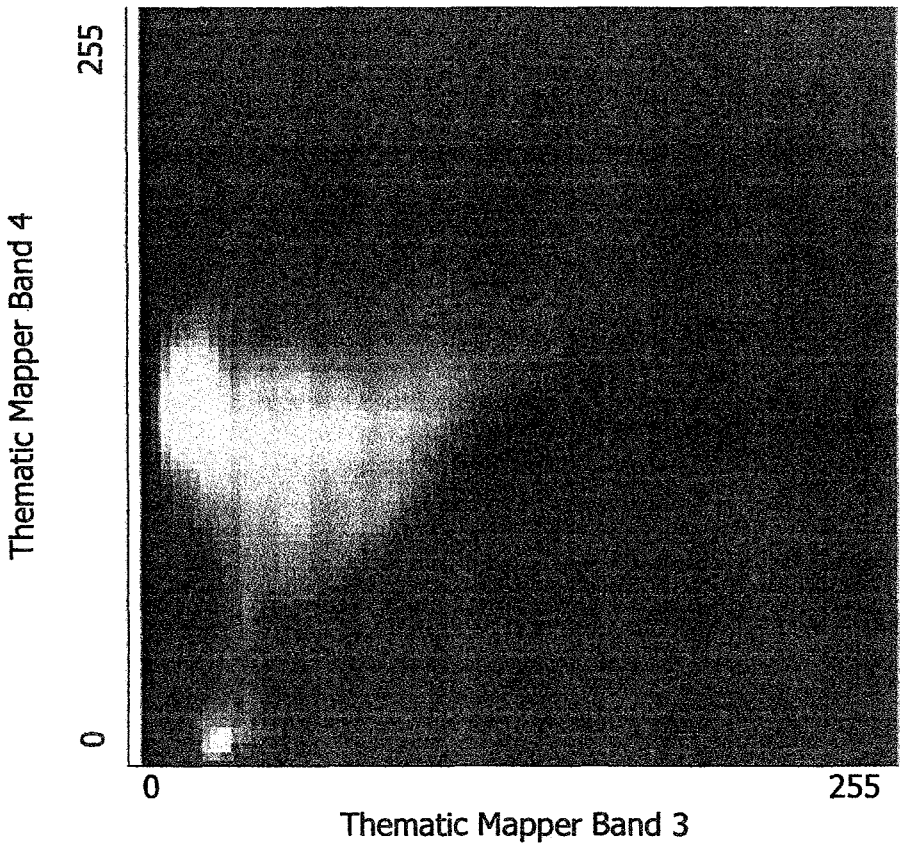


Figure 5-5: Feature space plot of Landsat-TM band 3 and 4 of Non Din Deang, Buriram, Thailand, obtained on 2 February 1999.

5.2 Radiometric Correction

5.2.1 Source of image radiometry error.

Bernstein (1983) stated that three main sources of image radiometry errors are detectors, A/D converter, and atmospheric effects.

1. Detectors. Current and future remote sensors use detectors that convert the sensed radiance into a voltage or digital number. Multiple detectors are used in the sensor systems, and due to variation in individual detector outputs, there will be different output for the same ground radiances. It is apparent that a different detector gain or bias will cause the detector output to vary.

2. A/D Converter. An analog-to-digital converter can cause both quantization and signal errors. Because the converter has a finite numbers of bits, quantization noise with a mean square value of q2/12 will result, where q is the quantization value (least significant bit) of the digitized signal. Select a converter with a sufficient number of bits (8 bits or better) will minimize this effect.

If the A/D converter has a bias or gain error, it will also degrade the data in a manner similar to a detector gain or bias error.

3. Atmospheric Effects. The atmosphere introduces both scattering and attenuation effects. Although not an error, as this is the "signal", techniques have been developed to compensate the sensor data for these effects. This usually the removal of the first order effects by increasing the gain or contents in the shorter wavelength bands.

5.2.2 Type of Radiometric Correction

Schowengerdt (1997) categorized radiometric correction in three levels included sensor calibration, atmospheric correction, solar and topographic correction as shown in Figure 5-6. The first converts the sensor digital number to at-sensor radiance and requires sensor calibration information. The second is transformation of the at-sensor radiance to radiance at the earth's surface. This level is much more difficult to achieve, since it require information about the atmospheric conditions at the time and location of the image. That information can be in different forms, ranging from a single categorization of the atmospheric condition as one of several "standard atmospheres" to estimates of certain parameters such as path radiance from the image data itself to ideally coincident ground measurement. A promising approach appears to be the combination of image-base data with atmospheric modeling. The final level of calibration to surface reflectance is achieved by correction for topographic slope and aspect, atmospheric path length variation due to topographic relief, particularly in mountainous terrain, solar spectral irradiance, solar path atmospheric transmittance, and down-scattered "skylight" radiance.

5.2.2.1 Sensor Calibration

The user of the data receives DN values and wants to convert them to radiance in an "inverse" calculation. Therefore, the user wants calibration coefficients with units of radiance-per-DN. These coefficients are often referred to as "gain" and offset. To calculate band-integrated radiance values at the sensor, we need to know the calibration gains and offsets of sensors as shown in Table 5-4 as an example for Landsat 4 and 5. Then these coefficients can be applied to pixel values in each band, DNb to produce band-integrated radiance values, L_b^s, , as follow:

$$L_b^s = gain_b \cdot DN_b + offset_b \tag{5-11}$$

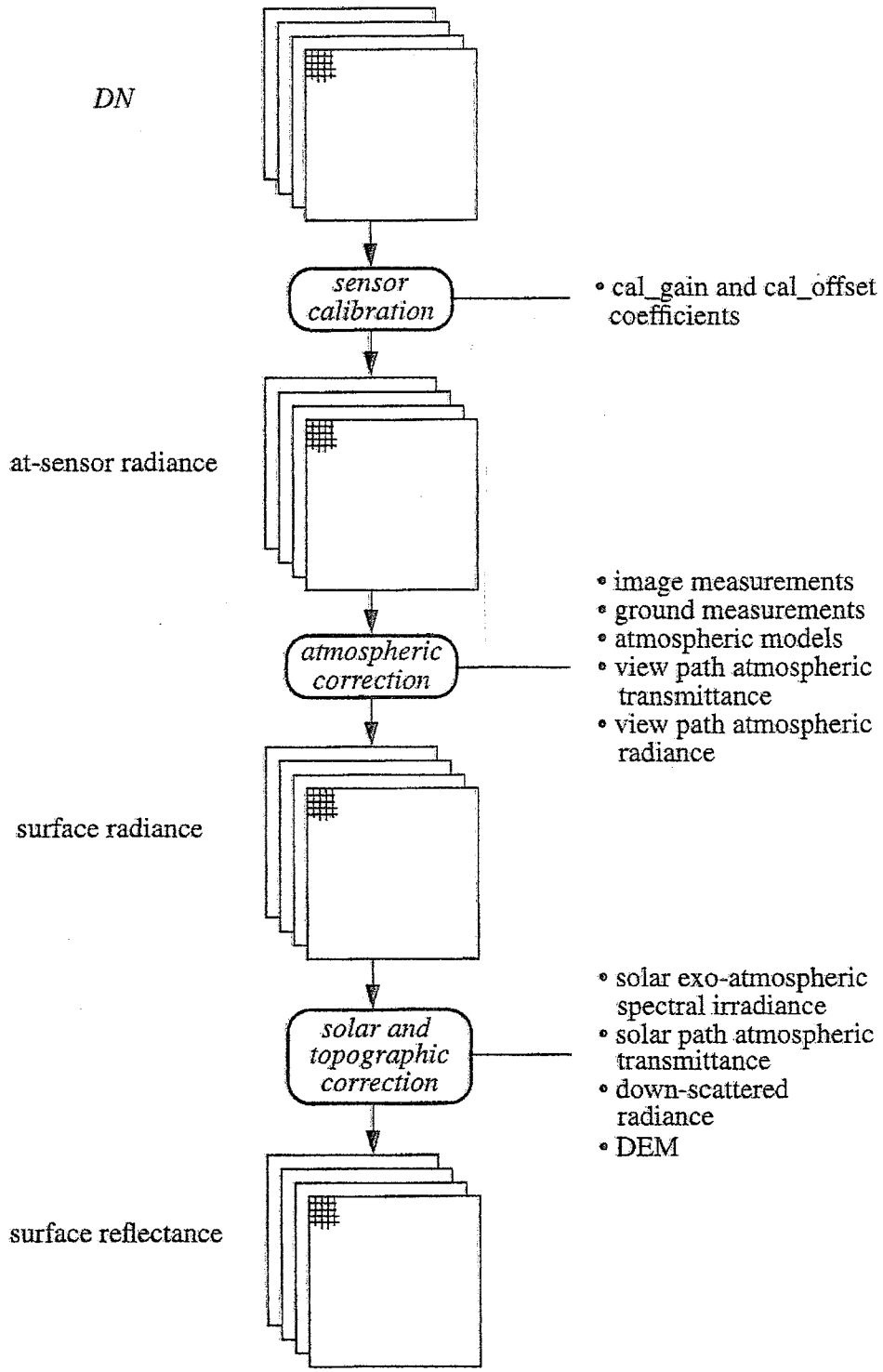


Figure 5-6: Data flow for calibration of remote sensing images to physical units (From Schowengerdt, 1997).

Table 5-4: Pre-launch measurement of the TM calibration gain and offset coefficient. Gain is given here in radiance/DN units and offset is radiance, with units $W\text{-}m^{-2}\text{-}sr^{-1}\text{-}\mu m^{-1}$

Band	Landsat-4		Landsat-5	
	Gain	Offset	Gain	Offset
1	0.672	-3.361	0.642	-2.568
2	1.217	-6.085	1.274	-5.098
3	0.819	-4.917	0.979	-3.914
4	0.994	-9.936	0.925	-4.629
5	0.120	-0.7208	0.127	-0.763
6	0.0568	+1.252	0.0552	+1.238
7	0.0734	-0.367	0.0677	-0.0338

From Schowengerdt, 1997.

Furthermore, Lillesand et al. (2004) explained that normally, detector and data system are designed to produce a linear response to incident spectral radiance. For example, Figure 5-7 shows the linear radiometric response function typical of an individual TM channel. L_{Min} is the spectral radiance corresponding to a DN response of 0 and L_{max} is the minimum radiance required to generate the maximum DN (here 255). Thus, L_{max} represents the radiance at which the channel saturated. The range from L_{Min} to L_{max} is the dynamic for the channel. Each spectral band of the sensor has its own response function, and its characteristics are monitored using onboard calibration lamps and temperature references for the thermal channel. The absolute spectral radiance output of the calibration sources is known from pre-launch calibration and is assumed to be stable over the life of the sensor. However, it can be modified over time as shown in Table 5-5. Thus, the onboard calibration sources from the basis for constructing the radiometric response function by relating known radiance values incident on the detectors to be resulting DNs.

The linear relationship between radiance and DN values for any given channel is yield as:

$$DN = GL + B \quad (5\text{-}12)$$

where

DN is digital number valued recorded

G is slope of response function (channel gain)

L is spectral radiance measured over the spectral bandwidth of the channel

B is intercept of response function (channel offset)

If we plot the inverse of the radiometric response (Figure 5-8), the equation for this line can be calculated as:

$$L = \left(\frac{L_{Max} - L_{Min}}{255} \right) DN + L_{Min} \quad (5\text{-}13)$$

Equation 5-13 can be used to convert any DN in a particular band to absolute units of spectral radiance in the band if LMAX and LMIN are known from the sensor calibration as same as Equation 5-11. Often the LMAX and LMIN values are expressed in unit W-cm^{-2}-sr^{-1}-μm^{-1}.

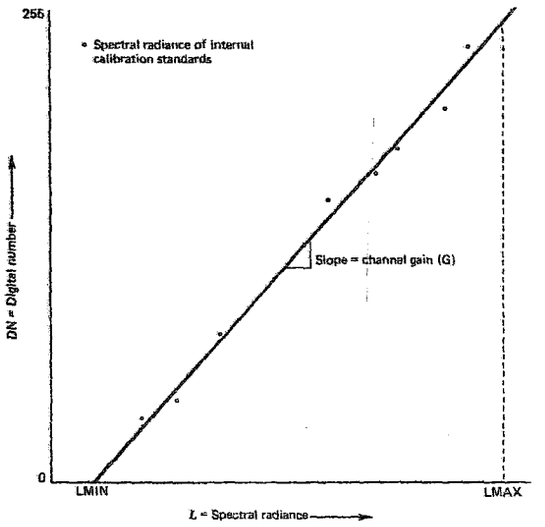


Figure 5-7: Radiometric response function for an individual TM channel (From Lillesand et al. 2004).

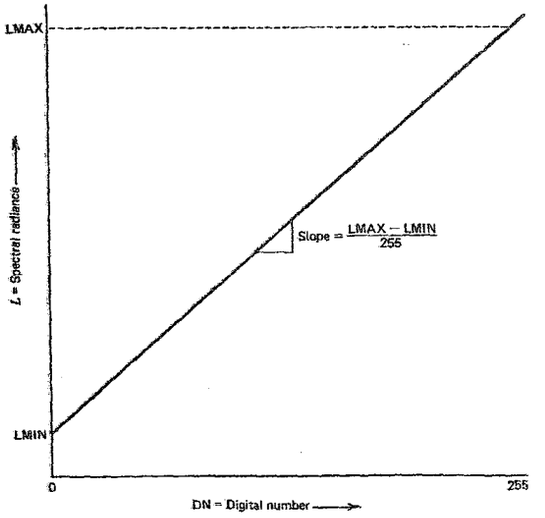


Figure 5-8: Inverse of radiometric response function for an individual TM channel. (From Lillesand et al. 2004).

Table 5-5: Value of L_{Min} and L_{Max} for the Landsat MSS sensor system on various platforms

Satellite	Spectral bands							
	4		5		6		7	
	L_{Min}	L_{Max}	L_{Min}	L_{Max}	L_{Min}	L_{Max}	L_{Min}	L_{Max}
Landsat 1	0.0	24.8	0.0	20.0	0.0	17.6	0.0	40.0
Landsat 2 (6/22 to 7/16/75)	1.0	21.0	0.7	15.6	0.7	14.0	1.4	41.5
Landsat 2 (after 7/16/75)	0.8	26.3	0.6	17.6	0.6	15.2	1.1	39.1
Landsat 3 (3/5 to 6/1/78)	0.4	22.0	0.3	17.5	0.3	14.5	0.3	44.1
Landsat 3 (after 6/1/78)	0.4	25.9	0.3	17.9	0.3	14.9	0.3	38.3
Landsat 4	0.2	23.0	0.4	18.0	0.4	13.0	1.0	40.0

From Jensen, 2005.

5.2.2.2 Atmospheric Correction

Schowengerdt (1997) stated that the atmosphere plays a complex role in optical remote sensing. The atmospheric dramatically alters the spectral nature of the radiation reaching the sensor. It is impossible to apply the complete model to every pixel in a typical remote sensing image. Independent atmospheric data are almost never available for any given image, and certainly not over the full FOV. A robust approach to the problem is to estimate atmospheric parameters from the image itself. Any atmospheric correction that relies on physical modeling of the atmosphere, first requires correction of the sensor DNs to radiance at the sensor calibration.

He mentioned that the "bootstrap" approach has been the most common type of atmosphere correction since the early days of Landsat (Table 5-6). A usual simplification is to assume that the down-scattered atmospheric radiance, E_{λ}^{s}, in Total solar radiant at-sensor,(L_{λ}^{e}) In following equation is zero. So that, in band-integrated form,

at-sensor: $$L_b^s(x,y) \cong L_b^{su}(x,y) + L_b^{sp} \quad (5\text{-}14)$$

Assuming this equation is an equity and view path transmittance ($\tau_v(\lambda)$, we can rewrite Eq. (5-14) as,

At-sensor: $$L_b^s(x,y) = \tau_{vb} L_b(x,y) + L_b^{sp} \quad (5\text{-}15)$$

And solve for the surface radiance, L_b, in each band,

Earth's surface $$L_b(x,y) = \frac{L_b^S(x,y) - L_b^{sp}}{\tau_{vb}} \quad (5\text{-}16)$$

In the simplest corrections for the visible wavelength region, the effort is focused on the estimating the upwelling atmospheric path radiance, L_λ^{sp}, with the view path transmittance assumed to be one, or at least a constant over wavelength. To a first approximation, this emphasis is not unreasonable, since the path radiance is the dominant atmospheric effect in the visible spectral region. The approach most commonly used requires identification of a "dark object' within the image, estimation of the signal level over the object, and subtraction of that level from every pixel in the image. The rationale is that the only possible source for an observed signal from a dark object is the path radiance. The major weaknesses in this approach are in identifying a suitable object in a given scene, and once one is located, in the assumption of zero reflectance for that object. In addition, the "dark object subtraction (DOS) technique cannot correct for the path transmittance term (τ_{vb}) in Equation 5-16, which is especially important in TM band 4, 5 and 7.

ERDAS (2002) stated that the dark pixel subtraction technique assumes that the pixel of lowest DN in each band should really be zero, and hence its radiometric value (DN) is the result of atmosphere-induced additive errors. These assumptions are very tenuous and recent work indicates that this method may actually degrade rather than improve the data.

Jensen (2005) stated that there are several ways to atmospherically correct remotely sensed data. Some are relatively straightforward while others are complex, being founded on physical principles and requiring a significant amount of information to function properly. He categorized major types of atmospheric correction including (1) absolute atmospheric correction, and (2) relative atmospheric correction. The details of these types will briefly be explained and demonstrated based on Jensen (2005 in the next section.

A. Absolute Atmospheric Correction

In general, solar radiation is largely unaffected as it travels through the vacuum of space. When it interacts with the Earth's atmosphere, however, it is selectively scattered and absorbed. The sum of these two forms of energy loss is called *atmospheric attenuation.* Atmospheric attenuation may (1) make it difficult to relate hand-held *in situ* spectroradiometer measurements with remote measurements, (2) make it difficult to extend spectral signatures through space and time, and (3) have an impact on classification accuracy within a scene if atmospheric attenuation varies significantly throughout the image.

Table 5-6: Example of atmospheric calibration techniques.

Sensor	Approach	Reference
MSS	Band-to-band regression	Potter and Mendolowitz, 1975
MSS	All band spectral covariance	Switzer et al., 1981
Airborne MSS	Band-to-band regression	Potter, 1984
AVHRR	Iterative estimation	Singh and Cracknell, 1986
MSS, TM	DOS with exponential scattering model	Chavez, 1988
TM	DOS with exponential scattering model Downwelling atmospheric radiance measurement	Chavez, 1989
TM	Pixel-by-pixel tasseled cap haze parameter	Lavreau, 1991
AVHRR	DOS, NDVI, AVHRR band 3	Holben et al., 1992
Airborne TMS, Landsat TM	Ground and airborne solar measure mean, atmospheric modeling code	Wrigley et al., 1992
TM	Comparison of ten DOS and atmospheric modeling code variations with field data	Moran et al, 1992
TM	Dark target, modeling code	Teilet and Fedosejevs, 1995
TM (all bands)	Atmospheric modeling code, region histogram matching	Richter, 1996a, Richter, 1996b
TM	DOS with estimated atmospheric transmittance	Chavez, 1996

From Schowengerdt, 1997.

The general goal of absolute radiometric correction is to turn the digital brightness values recorded by a remote sensing system into scaled surface reflectance values. These values can then be compared or used in conjunction with scaled surface reflectance values obtained anywhere else on the planet.

In practice, atmospheric radiative transfer codes (models) are firstly used to estimate the effects of atmospheric scattering and absorption effecting on satellite imagery. Then, each band and/or pixel in the scene can be adjusted to remove the effects of scattering and/or absorption. However, the application of these codes to a specific scene and date also requires knowledge of both the sensor spectral profile and the atmospheric properties at the same time. Target and path

radiance is one of important atmospheric properties associated with absolute atmospheric correction.

Target and Path Radiance

Ideally, the radiance (L) recorded by the camera or detector is a true function of the amount of radiance leaving the target terrain within the instantaneous-field-of-view (IFOV) at the specific solid angle. Unfortunately, other radiant energy may enter into the filed of view from various other paths and introduce confounding noise into the remote sensing process. Therefore, additional radiometric variable definitions are needed to identify the major sources and paths of this energy (Jensen, 2005). The variables were summarized by Jensen (2005) in Table 5-7. The various paths and factors that determine the radiance reaching the remote sensor are illustrated in Figure 5-9 including:

- Path 1 contains spectral solar irradiance that was attenuated very little before illuminating the terrain within the IFOV. Notice in this case that we are interested in the solar irradiance from a specific solar zenith angle and that the amount of irradiance reaching the terrain is a function of the atmospheric transmittance at this angle. If all of the irradiance makes it to the ground, then the atmospheric transmittance equals one. If none of the irradiance makes it to the ground, then the atmospheric transmittance is zero.
- Path 2 contains spectral diffuse sky irradiance that never even reaches the Earth's surface (the target study area) because of scattering in the atmosphere. Unfortunately, such energy is often scattered directly into the IFOV of the sensor system. It contains much unwanted diffuse sky irradiance that was inadvertently scattered into the IFOV of the sensor system. Therefore, if possible, we want to minimize its effects. Green (2003) refers to the quantity as the upward reflectance of the atmosphere.
- Path 3 contains energy from the Sun that has undergone some Rayleigh, Mie, and/or nonselective scattering and perhaps some absorption and reemission before illuminating the study area. Thus, its spectral composition and polarization may be somewhat different from the energy that reaches the ground from path 1. Green (2003) refers to this quantity as the downward reflectance of the atmosphere.
- Path 4 contains radiation that was reflected or scattered by nearby terrain covered by snow, concrete, soil, water, and/or vegetation into the IFOV of the sensor system. The energy does not actually illuminate the study area of interest. Therefore, if possible, we would like to minimize its effects.

- Path 5 is energy that was also reflected from nearby terrain into the atmosphere, but then scattered or reflected onto the study area.

Image offset (Deskew) caused by earth rotation effect

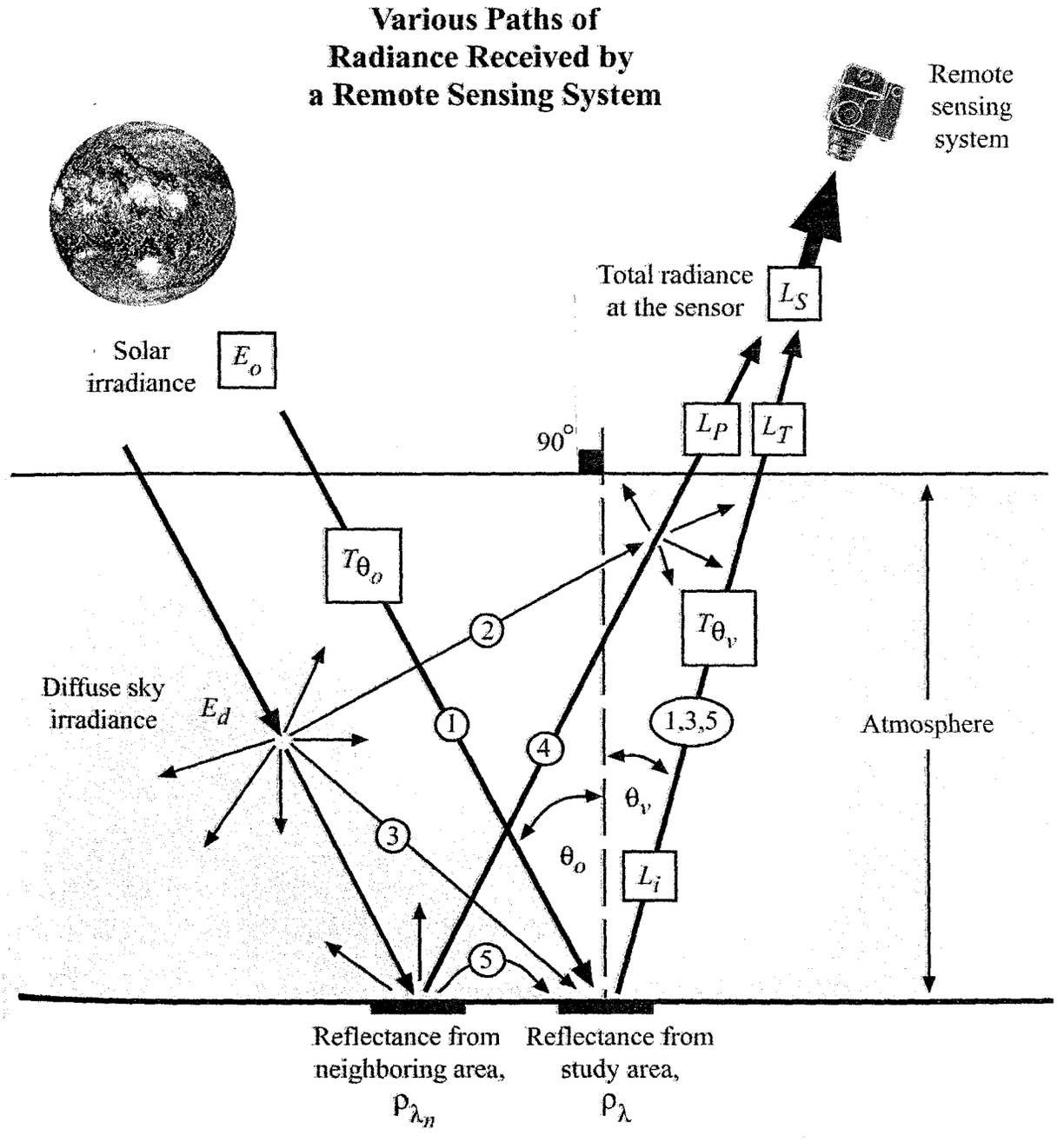


Figure 5-9: Various paths of radiance received by a remote sensing system (Jensen, 2005)

Table 5-7: Radiometric variables used in remote sensing

Symbol	Radiometric variable
E_o	solar irradiance at the top of the atmosphere (W m^{-2})
E_{o_λ}	spectral solar irradiance at the top of the atmosphere (W m^{-2} μm^{-1})
E_d	diffuse sky irradiance (W m^{-2})
E_{d_λ}	spectral diffuse sky irradiance (W m^{-2} μm^{-1})
E_{du_λ}	the upward reflectance of the atmosphere
E_{dd_λ}	the downward reflectance of the atmosphere
E_g	global irradiance incident on the surface (W m^{-2})
E_{g_λ}	spectral global irradiance incident on the surface (W m^{-2} μm^{-1})
τ	normal atmospheric optical thickness
T_θ	atmospheric transmittance at the angle θ to the zenith
θ_o	solar zenith angle
θ_v	View angle of the satellite sensor (or scan angle)
μ	Cos θ
ρ_λ	surface target reflectance at the specific wavelength
$\rho_{\lambda n}$	reflectance from a neighboring area
L_s	Total radiance at the sensor (W m^{-2} sr^{-1})
L_T	Total radiance from target of interest toward the sensor (W m^{-2} sr^{-1})
L_I	intrinsic radiance of the target (W m^{-2} sr^{-1})
L_P	Path radiance from multiple scattering (W m^{-2} sr^{-1})

From Jensen, 2005.

Therefore, for a given spectral interval in the electromagnetic spectrum (e.g. λ_1 to λ_2 could be 0.6 – 0.7 μm or red light), the total solar irradiance reaching the Earth's surface, E_{g_λ}, is an integration of several components:

$$E_{g_\lambda} = \int_{\lambda_1}^{\lambda_2} (E_{o_\lambda} T_{\theta_o} \cos\theta_o + E_{d_\lambda}) d\lambda \quad (5\text{-}17)$$

Only a small amount of this irradiance is actually reflected by the terrain in the direction of satellite sensor system. If we assume the surface of Earth is a diffuse reflector (a Lambertian surface), the total amount of radiance existing the target study area (L_τ) toward the sensor is:

$$L_T = \frac{1}{\pi} \int_{\lambda_1}^{\lambda_2} \rho_\lambda T_{\theta v} (E_{o_\lambda} T_{\theta_o} \cos\theta_o + E_{d_\lambda}) d\lambda \quad (5\text{-}18)$$

At the same time, some additional radiance from different paths that may fall within the IFOV of the sensor system detector. This is often called ***path radiance***, L_P. Thus, the total radiance reaching the sensor becomes:

$$L_S = L_T + L_P \quad (5\text{-}19)$$

In fact, the path radiance is an intrusive bad component of the total amount of radiances recorded by the sensor system. It is composed of radiant energy mainly from the diffuse sky irradiance (E_d) from path 2 as well as the reflectance from nearby ground area (ρ_{λ_n}) from path 4. A great deal of research has gone into develop methods to remove the contribution of path radiance. To understand how to remove atmospheric attenuation, it is useful to review the fundamental mechanisms of scattering and absorption related to atmospheric transmittance.

Atmospheric Transmittance

In the absence of atmosphere, the transmittance of solar radiant energy to the ground would be 100%. However, because of absorption and scattering, not all of the radiant energy reaches the ground. The amount that does reach the ground, relative to that no atmosphere, is called transmittance. Atmospheric transmittance (T_θ) may be computed as:

$$T_\theta = e^{-\tau / \cos\theta} \quad (5\text{-}20)$$

where τ is the normal atmospheric optical thickness and θ can represent either θ_o or θ_v, in Figure 5-9. The optical thickness of the atmosphere at certain wavelength, τ (λ), equals the sum of all the attenuation coefficients, which are made up primarily of Rayleigh scattering (τ_m) Mie scattering (τ_p) and selective atmospheric absorption (τ_a):

$$\tau(\lambda) = \tau_m + \tau_p + \tau_a \quad (5\text{-}21)$$

where

$$\tau_a = \tau_{H_2O} + \tau_{O_2} + \tau_{O_3} + \tau_{CO_2}$$

Basically, atmospheric scattering may add brightness to landscape spectral measurement while atmospheric absorption may subtract brightness from landscape spectral measurement.

Diffuse Sky Irradiance

Some additional scene irradiance comes from scattered skylight. Path 3 in Figure 5-9 contains radiation that has undergone some scattering before illumination the study area. Similarly, Path 5 contains energy that was reflected from a nearby area on the ground area on the ground into atmosphere and then scattered once again on the study area. The total diffuse sky irradiance at the pixel is E_d.

Hence, it is possible to determine how atmospheric effects (transmittance, diffuse sky irradiance, and path irradiance) affect the radiance by the remote sensing system. First, however, because the bandwidth used in remote sensing are relatively narrow (e.g., 0.5-0.6 μm), it is possible to restate Equation 5-17 and 5-18 without the integral. The total irradiance at the Earth's surface can be rewritten as:

$$E_g = E_{o\Delta\lambda} T_{\theta_o} \cos\theta_o \Delta\lambda + E_d \quad (5\text{-}22)$$

where : $E_{\lambda\Delta o}$ is the average spectral irradiance in the band interval $\lambda_\Delta = \lambda_2 - \lambda_1$. The total radiance transmitted through the atmosphere toward the sensor (LT) become:

$$L_T = \frac{1}{\pi} \rho T_{\theta_v} (E_{o\Delta\lambda} T_{\theta_o} \cos\theta_o \Delta\lambda + E_d) \quad (5\text{-}23)$$

The total radiance reaching the sensor then becomes:

$$L_S = \frac{1}{\pi} \rho T_{\theta_v} (E_{o\Delta\lambda} T_{\theta_o} \cos\theta_o \Delta\lambda + E_d) + L_P \quad (5\text{-}24)$$

This total radiance is used to relate brightness values in remotely sensed data to measured radiance, using the equation:

$$L_S = (K \times BV_{i,j,k}) + L_{\min} \quad (5\text{-}25)$$

Where

K is radiance per bit of sensor count rate $[(L_{Max} - L_{Min})/C_{Max}]$

$BV_{i,j,k}$ is brightness value of a pixel

C_{Max} is maximum value in the dataset (e.g., 8-bit = 255)

L_{Max} is radiance measured at detector saturation (W m^{-2} sr^{-1})

L_{Min} is lowest radiance measured at detector saturation (W m^{-2} sr^{-1})

A.1 Absolute Atmospheric Correction Based on Radiative Transfer Modeling

ERDAS (2002) stated that atmospheric modeling is computationally complex and requires either assumptions or inputs concerning the atmosphere at the time of imaging. The atmospheric model used to define the computations is frequently Lowtran or Modtran.

Jensen (2005) mentioned that most current radiative transfer-based atmospheric correction algorithm can compute and characterize the scattering and absorption characteristics of the atmosphere on a specific date and times. However, the users need to provide information include:

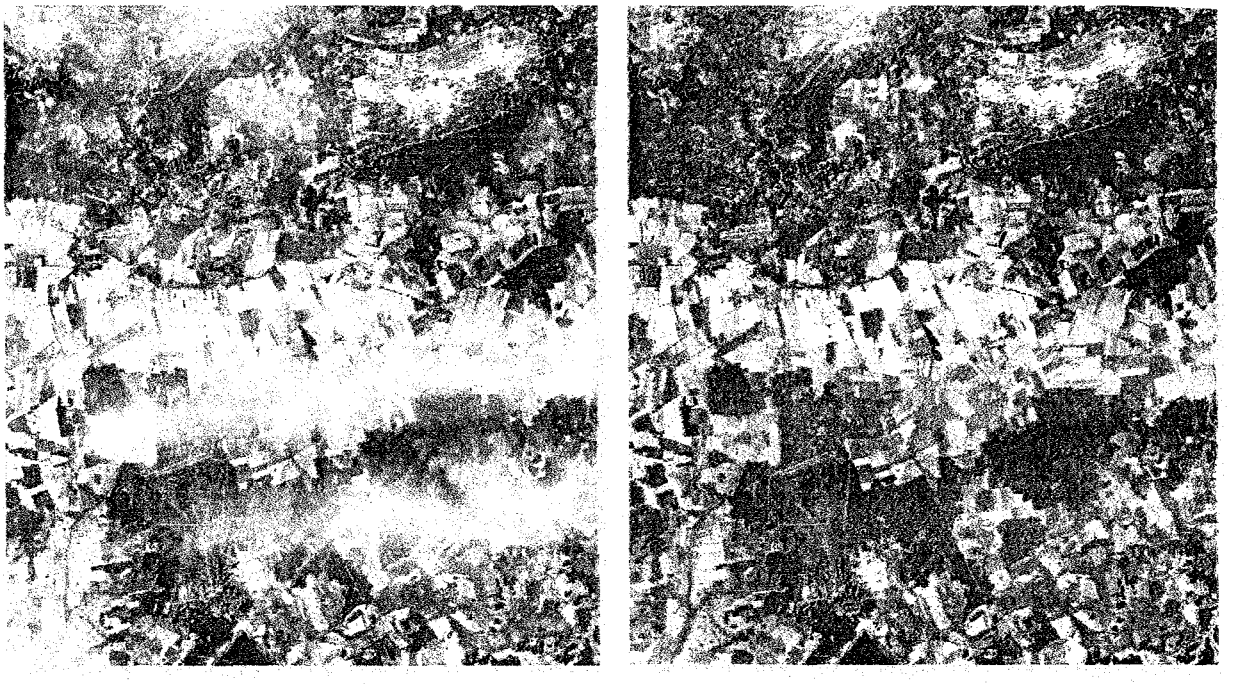
- latitude and longitude of the remotely sensed image scene,
- date and exact time of remote sensing data collection,
- image acquisition altitude,
- mean elevation of the scene,
- an atmospheric model (e.g., mid-latitude summer, mid-latitude winter),
- radiometrically calibrated image radiance data,
- data about each specific band (i.e., its mean and full-width at half-maximum,
- local atmospheric visibility at the time of remote sensing data collection.

Numerous atmospheric correction algorithms have been developed based on the use of radiative transfer principle. Several of the most important include:

- ACORN- **A**tmospheric **COR**rection **N**ow,
- ATREM- **AT**mospheric **REM**oval program,
- FLAASH- **F**ast **L**ine-of-sign **A**tmospheric **A**nalysis of **S**pectral **H**ypercubes,
- ATCOR- **AT**mospheric **COR**rection. (See example in Figure 5-10)

ERDAS (2002) mentioned that accurate atmospheric modeling is essential in preprocessing hyperspectral data sets where bandwidths are typically 10 nm or less. These narrow bandwidth

corrections can then be combined to simulate the much wider bandwidths of Landsat or SPOT sensors.



a. Before atmospheric correction. b. After atmospheric correction.

Figure 5-10: An example absolute atmospheric correction using ATCOR (From Jensen, 2005)

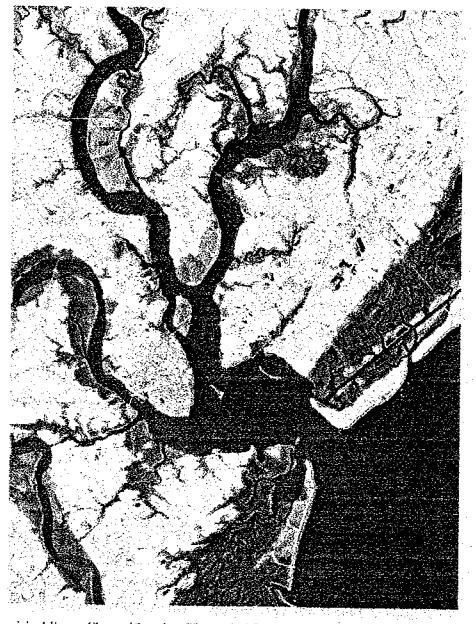
A.2 Absolute Atmospheric Correction Using Empirical Line Calibration

Absolute atmospheric correction may also be performed using **empirical line calibration (ELC)**, which forces the remote sensing image data to match ***in situ*** spectral reflectance measurements, hopefully obtained at approximately the same time and on the same date as the remote sensing overflight (Jensen, 2005). Empirical line calibration is based on the equation:

$$BV_k = \rho_\lambda A_k + B_k \qquad (5\text{-}26)$$

where BVk is the digital output value for a pixel in band k, p_λ equals the scaled surface reflectance of the materials within the remote sensor IFOV at a specific wavelength (λ), A_k is a multiplicative term affecting the BV, and B_k is an additive term. The multiplicative term is associated primarily with atmospheric transmittance and instrumental factors, and the additive term deals primarily with atmospheric path radiance and instrumental offset (i.e., dark current).

To use empirical line calibration (ELC), the analyst usually selects two or more areas in the scene with different albedos (e.g., one bright target such as a sand pile and one dark target such as a deep, non-turbid water body). The areas should be as homogeneous as possible. *In situ* spectroradiometer measurements of these targets are made on the ground. The in situ and remote sensing derived spectra are regressed and gain and offset value computed. The gain and offset values are then applied to remote sensor data on a band by band basis, removing atmospheric attenuation. An example of empirical line calibration (ELC) was shown in Figure 5-11.



a. Empirical line calibrated Landsat Thematic Mapper band 4 image of Charleston, SC.

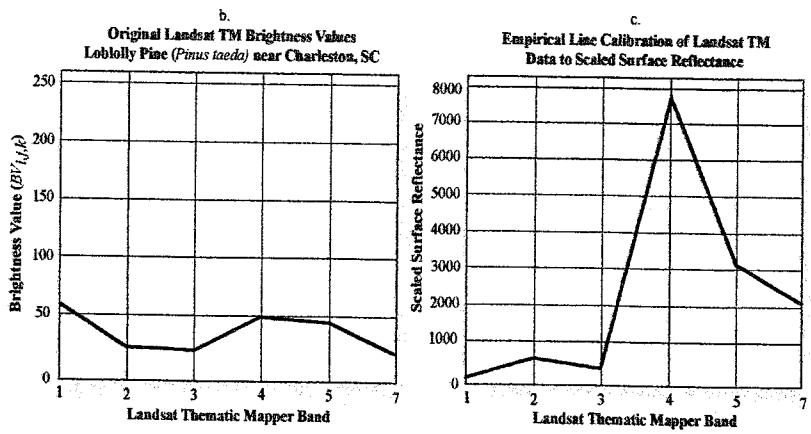


Figure 5-11: Absolute Atmospheric Correction Using Empirical Line Calibration (From Jensen, 2005)

B. Relative Atmospheric Correction

The application of atmospheric radiative transfer codes (models) to a specific scene and date also requires knowledge of both the sensor spectral profile and the atmospheric properties at the same time of remote sensor data collection. If all this information is available, then it is possible for the atmospheric radiative transfer code to provide realistic estimate of the effects of atmospheric scattering and absorption and convert the imagery to scaled surfaced reflectance. However, atmospheric properties are difficult to acquire even when planned for. For this reason, relative radiometric correction techniques have been developed (Jensen, 2005). Relative radiometric correction may be used:

- to normalized the intensities among the different bands within a single date remotely sensed image
- to normalize the intensities of bands of remote sensor data in multiple dates of imagery to a standard scene selected by the analyst.

B.1 Single-image Normalization Using Histogram Adjustment

This simple method is based on the fact that infrared data (> 0.7 μm) are largely free of atmospheric scattering effects, whereas the visible region (0.4-0.7 μm) is influenced by them. The method involves evaluating the histograms of the various bands of remotely sensed data of the desired scene. Normally, the data collected in the visible wavelengths (e.g., TM band 1 to 3) have high minimum value because of the increased atmospheric scattering taking place in these wavelengths. Conversely, atmospheric absorption subtracts brightness from the data recorded in the longer-wavelength interval (e.g., TM band 4, 5 and 7). This effect commonly causes data from the infrared bands to have minimums close to zero, even when few objects in the scene truly have a reflectance of zero (Figure 5-12a).

If the histograms are shifted to the left so that zero values appear in the data, the effects of atmospheric scattering will be minimized. This simple algorithm models the first-order of atmospheric scattering, or haze. It is based on subtractive bias established for each spectral band. The bias may determined by evaluating of a histogram of brightness values of a reference target such as deep water in all bands. The atmospheric effects correction algorithm is defined:

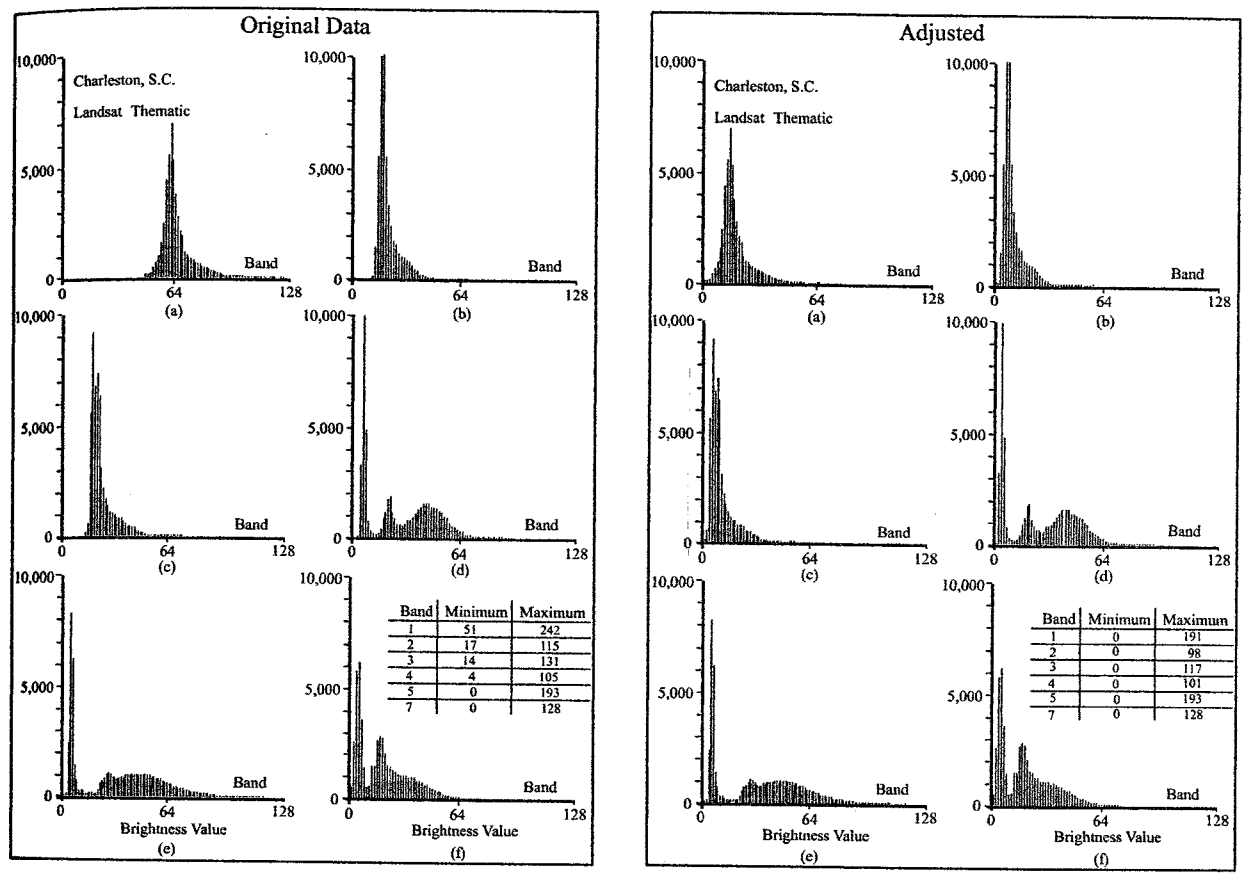
$$\text{Output } BV_{i,j,k} = \text{Input } BV_{i,j,k} - \text{Bias} \qquad (5\text{-}27)$$

Where:

$\text{Input } BV_{i,j,k}$ is input pixel value at line i and column j of band k

$\text{Output } BV_{i,j,k}$ is the adjusted pixel value at same location

Figure 5-12b shows the result of applying a simple histogram adjustment atmospheric scattering correction.



(a) (b)

Figure 5-12: Single-image Normalization Using Histogram Adjustment (From Jensen, 2005)

B.2 Multiple-date Image Normalization Using Regression

Multiple-date image normalization involves selecting a base image (b) and then transform the spectral characteristics of all other images obtained on different date (e.g., those obtained on date b-1, b-2 and/or b+1, b+2, etc.) to have approximately the same radiometric scale as the base image. It is important to remember, however, that the radiometric scale used in relative multiple-date image normalization will be simple brightness value rather than scaled surface reflectance produced when conducting an absolute radiometric correction (Jensen, 2005).

ERDAS (2002) mentioned that a number of methods using linear regressions have been tried. These techniques use bispectral plots and assume that the position of any pixel along that plot is strictly a result of illumination. The slope then equals the relative reflectivities for the two spectral bands. At an illumination of zero, the regression plots should pass through the bispectral origin. Offsets from this represent the additive extraneous components, due to atmosphere effects.

Multiple-date image normalization involves selection of *pseudo-invariant feature* (PIFs), often referred to as radiometric ground control point (Jensen, 2005). The selected pseudo-invariant feature should have the following characteristics:

- The spectral characteristics of PIF should change very little through time. Deep non-turbid water bodies, bare soil, large rooftops, or other homogeneous feature are candidates.
- The PIF should be approximately the same elevation as the other land in the scene.
- The PIF should normally contain only minimal amounts of vegetation. Vegetation spectral reflectance can change overtime as a result of environmental stress and plant phenology. However, stable, homogeneous forest canopy image might be considered.
- The PIF must be in a relatively flat area so that incremental changes in Sun angle from date to date will have the same proportional increase or decrease in direct beam sunlight for all normalization target.

Regression is used to relate the base image's PIF spectral characteristics with PIF spectral characteristics from other dates. The algorithm assumes that the pixels sampled at time b+1 or b-1 are linearly related to the pixel for the same locations on the base image (b). This implied that the spectral reflectance properties of the sampled pixels have not change during the time interval. Thus, the key to image regression method is the selection of quality pseudo-invariant feature. The result of normalization of multiple-date using regression was shown in Figure 5-13.

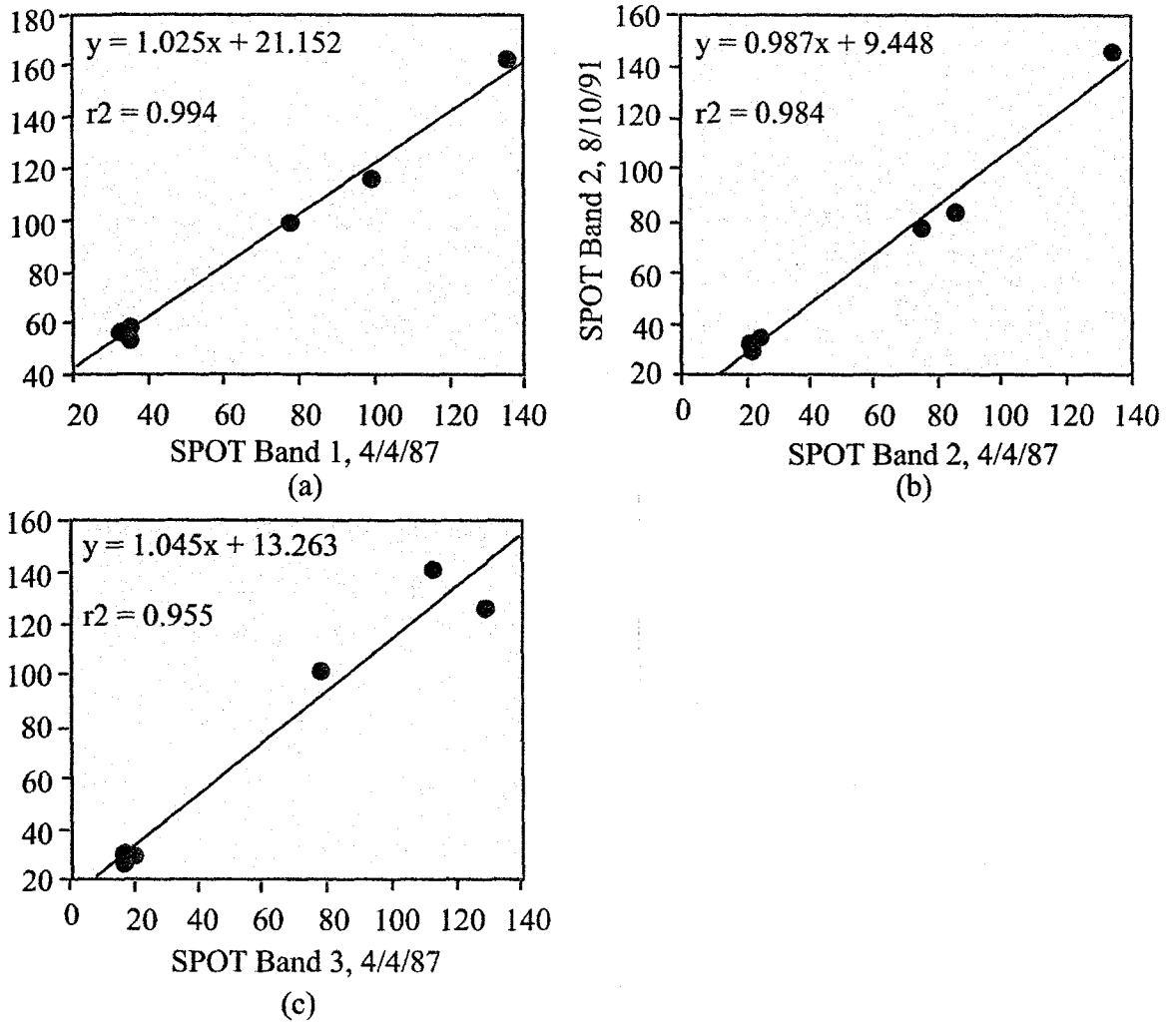


Figure 5-13: Multiple-date Image Normalization Using Regression (From, Jensen, 2005)

5.2.2.3 Solar and Topographic Correction

Schowengerdt (1997) stated that to find the spectral reflectance of every pixel in an image required more external information and further correction by using total solar radiant at-sensor and solving Equation 5-16 for reflectance in terms of the surface reflectance radiance:

$$\rho_b(x,y) = \frac{\pi L_b(x,y)}{\tau_{sb} E_b^0 \cos[\theta(x,y)]} \quad (5\text{-}28)$$

Thus, we need the surface radiance (Eq. 5-16), solar path atmospheric transmittance, exo-atmospheric solar spectral irradiance and the incident angle which depends on the solar angles and topography. Again, in the visible, the atmospheric transmittance is reasonably predictable for

different atmospheric conditions. Ideally, however, it should be measured at the time of image acquisition. The exo-atmospheric solar spectral irradiance is a well known quantity and for flat terrain, the incident angle is readily obtained. For non-flat terrain, calculation of incident angle at each pixel requires a DEM at the same Ground Sample Interval (GSI) the sensor image.

Jensen (2005) stated that topographic slope and aspect may also introduce in the radiometric distortion of the recorded signal. In some locations, the area of interest might even be in complete shadow, dramatically affecting the brightness values of the pixels involved. For these reason, research has been directed toward the removal of topographic effects, especially in mountainous regions on multi-spectral data. The goal of a slope-aspect correction is to remove topographically induced illumination variation so that two objects having the same reflectance properties show the same brightness value in the image despite their different orientation to the Sun's position. Teillet et al. (1982) described four topographic slope-aspect correction methods included (1) the simple cosine correction, (2) the Minnaert correction, (3) the C correction and a statistic-empirical correction. Each correction is based on illumination, which is defined as the cosine of the incident solar angle, thus representing the proportion of the direct solar radiation hitting a pixel. The amount of illumination is dependent on the relative orientation of the pixel toward the Sun's actual position (Figure 5-14). Each slope-aspect topographic correction method requires a DEM of the study area. The DEM is processed so that each pixel's brightness value represents the amount of illumination it should receive from the Sun. This information is then modeled using one of four algorithms to enhance or subdue the original brightness value of remote sensor data.

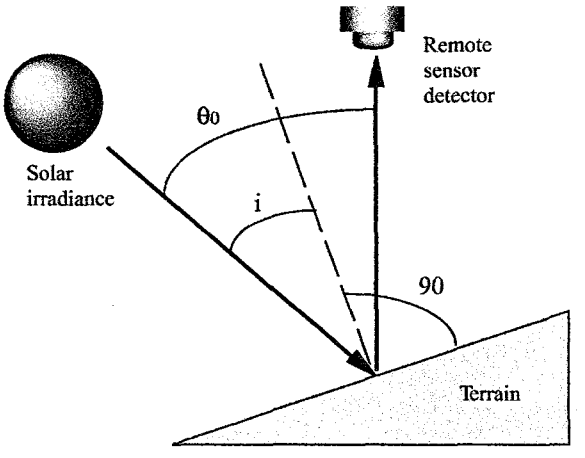


Figure 5-14: Representation of the Sun's angle of the incidence (i) and the solar zenith angle (θ_o) (From Jensen, 2005).

A. The Cosine Correction

The amount of irradiance reaching a pixel on a slope is directly proportional to the cosine of the incidence angle i, which is defined as the angle between the normal on the pixel in question and zenith direction (Teillet et al., 1982). It assumes that:

- Lambertian surface,
- a constant distance between Earth and the Sun, and
- constant amount of solar energy illuminating Earth.

Therefore, only the part of cos i of the total incoming irradiance, Eg, reaches the inclined pixel. Thus, it is possible to perform a simple topographic slope-aspect correction of remotely sensed data using the following cosine equation:

$$L_H = L_T \frac{\cos\theta_o}{\cos i} \quad (5\text{-}29)$$

where:

LH is radiance observed for a horizontal surface (corrected data);

LT is radiance observed over slope terrain (raw data);

θo is Sun's zenith angle;

i is Sun's incidence angle in relation to the normal on a pixel (Figure 5-14)

B. The Minnaert Correction

Teillet et al. (1982) introduced the Minaert correction to the basic cosine function:

$$L_H = L_T \left(\frac{\cos\theta_o}{\cos i} \right)^k \quad (5\text{-}30)$$

where:

k is the Minnaert constant

The constant varies between 0 and 1 and is a measure of the extent to which a surface is Lambertian. A perfect Lambertian surface has k = 1 and represents a traditional cosine correction.

C. The C Correction

Teillet et al.(1982) introduced an additional adjustment to the cosine function called the c correction:

$$L_H = L_T \frac{\cos\theta_o + c}{\cos i + c} \quad (5\text{-}31)$$

where c is the division of y-intercept of regression line (b) by slope of the regression line (m).

D. A Statistical-Empirical Correction

For each pixel in the scene, it is possible to correlate the predicted illumination (cos i x 100) from the DEM with the actual remote sensor data. For example, Meyer et al. (1993) correlated Landsat-TM data of known forest stand in Switzerland with the predicted illumination from high-resolution DEM. Any slope in the regression line suggests that a constant type of forest stand will appear differently on different terrain slopes. Conversely, by taking into account the statistical relationship in the distribution, the regression line may be rotated on the following equation (Jensen, 2005):

$$L_H = L_T - \cos(i)m - b + \overline{L_T} \qquad (5\text{-}32)$$

where:

LH is radiance observed for a horizontal surface (corrected data);

LT is radiance observed over slope terrain (raw data);

$\overline{L_T}$ is average of L_T form forest pixels (according to ground reference data)

i is Sun's incidence angle in relation to the normal on a pixel (Figure 5-14)

m is slope of the regression;

b is y-intercept of regression line

5.3 Geometric Correction

5.3.1 Source of Image Geometry Error.

Bernstein (1983) stated that major sources of image geometry errors included: altitude, attitude, scan-skew, velocity, earth rotation, map projection, sensor mirror sweep, panorama, and perspective.

1. **Altitude.** Departures of the spacecraft from nominal altitude produce scale distortions in the sensor data.

2. **Attitude.** Nominally, the sensor axis system is maintained with one axis normal to the Earth's surface and another parallel to the spacecraft velocity-vector. As the sensor departs from the attitude, geometric distortion results.

3. **Scan-Skew.** During the time required for the MSS mirror to complete an active scan, the spacecraft moves along the ground track. Thus, the ground swath scanned is not normal to the ground track but is slightly skewed, which produces cross-scan geometric distortion.

4. **Velocity.** If the spacecraft velocity vector departs from nominal values, the ground track covered by a given number of successive mirror sweeps changes, producing along-track scale distortion.

5. **Earth Rotation.** As the MSS mirror completes successive scans, the Earth rotates beneath the sensor. Thus, there is a gradual westward shift of the ground swath being scanned. This caused along-scan distortion.

6. **Map Projection.** For earth resource use, image data are usually required in a specific map projection. Although map projection does not constitute a geometric error, it does require a geometric transformation of the input data, and this can be accomplished by the same operations and compensate for distortion in the data.

7. **Sensor Mirror Sweep.** The MSS mirror scanning rate varies nonlinearly cross a scan because of imperfections in the electro-mechanical driving mechanism. Since data samples are taken at regulars of time, the varying scan rate produces along-scan distortion.

8. **Panorama.** The image ground-area is proportional to the tangent of the scan angle rather than to the angle itself and, since data samples are taken at regular intervals, this produces along-scan distortion.

9. **Perspective.** For most earth resources application desired Landsat images represent the projection of points on the Earth on the plane tangent to the Earth at the nadir, with all projection lines normal to the plane. For the MSS, this produces only along-scan distortion.

Basically, remotely sensed imagery typically exhibits ***internal and external geometric error.*** It is important to recognize the source of the internal and external error and whether it is *systematic* (predictable) or *nonsystematic* (random). Systematic geometric error is generally easier to identify and correct than random geometric error.

5.3.1.1 Internal Geometric Errors

Internal geometric errors are introduced by the remote sensing system itself or in combination with Earth rotation or curvature characteristics. These distortions are often systematic (predictable) and may be identified and corrected using pre-launch or in-flight platform ephemeris. Geometric distortions in imagery include:

- skew caused by Earth rotation effects,
- scanning system–induced variation in ground resolution cell size,
- scanning system one-dimensional relief displacement,
- scanning system tangential scale distortion.

A. Image Offset (skew) caused by Earth Rotation Effects

Earth-observing Sun-synchronous satellites are normally in fixed orbits that collect a path (or swath) of imagery as the satellite makes its way from the north to the south in descending mode. Meanwhile, the Earth below rotates on its axis from west to east making one complete revolution every 24 hours. This interaction between the fixed orbital path of the remote sensing system and the Earth's rotation on its axis skews the geometry of the imagery collected (Jensen,2005).

Deskewing is defined as the systematic displacement of pixels westward in a frame of imagery to correct for the interaction of the satellite sensor system's angular velocity and Earth's surface velocity. The amount of shift or displacement to the west is a function of the relative velocities of both the satellite and the Earth and the length of the image frame that is recorded. Most satellite data providers automatically deskew the data that is delivered to users using the logic shown in Figure 5-15.

Jensen (2005) explained the deskewing procedure in following steps:

Step I: The surface velocity of the Earth, vearth, is computed:

$$v_{earth} = \omega_{earth} r \cos\lambda \tag{5-33}$$

Where

r is the radius of the Earth (6.387816 Mm)

$\omega earth$ is the Earth's rotational velocity (72.72 μrad s^{-1} at a specific latitude, λ.

Step II: To determine the *length of time*, (*st*), it take for the satellite to scan a typical remote sensing frame (*L*) of data on the ground. For typical *185* km frame of Landsat- MSS imagery would be scanned in:

$$s_t = \frac{L}{r \times \omega_{land123}} \tag{5-34}$$

Where

r is the radius of the Earth (6.387816 Mm)

and $\omega_{land123}$ is an angular velocity of Landsat 1, 2, 3 with MSS of 1.014 mrad s^{-1}.

Step III: The surface of the Earth move to the east ($\Delta xeast$) by:

$$\Delta x_{east} = v_{earth} \bullet s_t \tag{5-35}$$

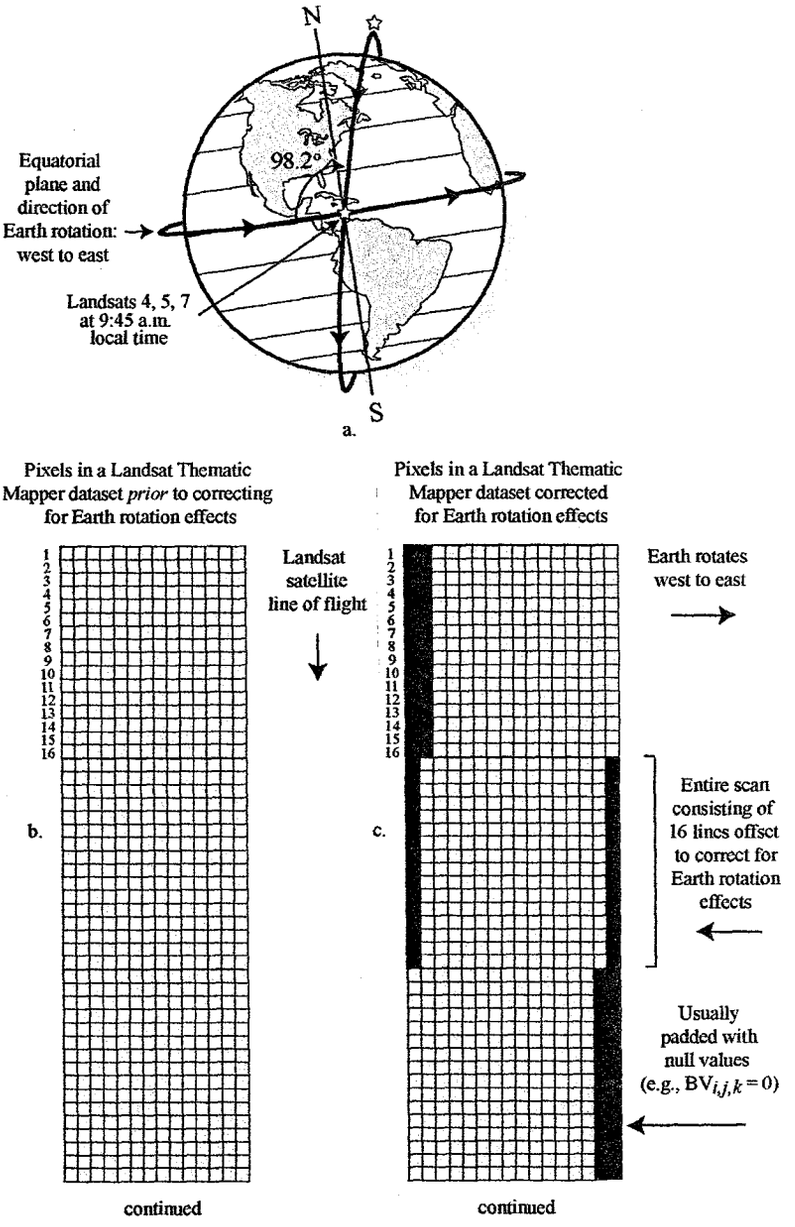


Figure 5-15 Image offset (Deskew) caused by earth rotation effect (a) Landsat satellites 4, 5, and 7 are in a Sun-synchronous orbit with an angle of inclination of 98.2°. The Earth rotates on its axis from west to east as imagery is collected. (b) Pixels in three hypothetical scans (consisting of 16 lines each) of Landsat TM data. While the matrix (raster) may look correct, it actually contains systematic geometric distortion caused by the angular velocity of the satellite in its descending orbital path in conjunction with the surface velocity of the Earth as it rotates on its axis while collecting a frame of imagery. (c) The result of adjusting (*deskewing*) the original Landsat TM data to the west to compensate for Earth rotation effects. Landsat 4, 5, and 7 use a bidirectional cross-track scanning mirror (From Jensen, 2005)

B. Scanning System-induced Variation in Ground Resolution Cell Size

An orbital multispectral scanning system scans through just a few degrees off-nadir as it collects data hundreds of kilometers above the Earth's surface (e.g., Landsat 7 data are collected at 705 km AGL). This configuration minimizes the amount of distortion introduced by the scanning system. Conversely, a suborbital multispectral scanning system may be operating just tens of kilometers AGL with a scan field of view of perhaps 70°. This introduces numerous types of geometric distortion that can be difficult to correct (Jensen, 2005). The configuration of scanning system-induced variation in ground resolution cell size geometrically displayed in Figure 5-16.

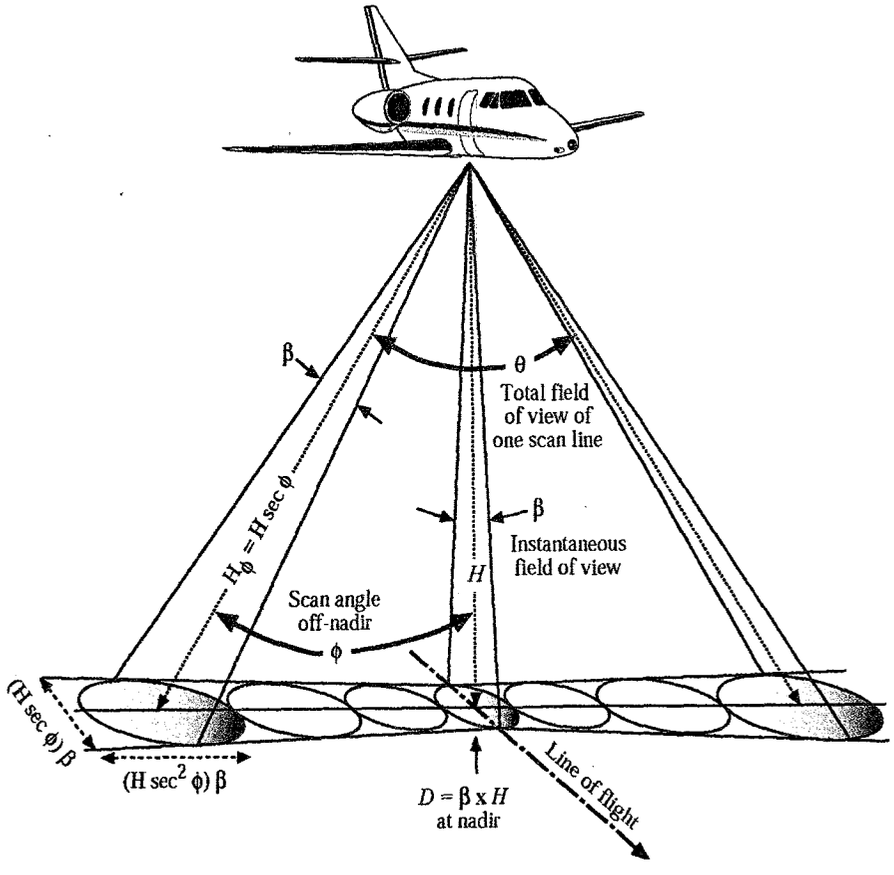


Figure 5-16: The geometrically configuration of scanning system-induced variation in ground resolution cell size (From Jensen, 2005).

B. 1 Ground Swath Width (gsw)

The ground swath width (gsw) is the length of the terrain strip remotely sensed by the system during one complete across-track sweep of the scanning mirror. It is a function of the total angular field of view of the sensor system, θ, and the altitude of the sensor system above ground level, H. It is computed as:

$$gsw = \tan\left(\frac{\theta}{2}\right) \times H \times 2 \quad (5\text{-}36)$$

B. 2 The diameter of the circular ground area viewed by the sensor, D

The diameter of the circular ground area viewed by the sensor, D, at nadir is a function of the instantaneous-field-of-view of the sensor,β, (mrad) and the altitude of the scanner above ground level (AGL), H, where D = β x H.

In fact, the distance from the aircraft to the resolution cell on the ground, HΦ, is a function of the scan angle off-nadir, Φ, at the time of data collection and the true altitude of the aircraft, H. It is computed as:

$$H_{\Phi} = H \times \sec\phi \quad (5\text{-}37)$$

Thus, the size of the ground-resolution cell increases as the angle increases away from nadir. The nominal (average) diameter of the elliptical resolution cell, DΦ, at this angular location from nadir has the dimension:

$$D_{\phi} = (H \times \sec\phi) \times \beta \quad (5\text{-}38)$$

In the direction of the line of flight, and

$$D_{\phi} = (H \times \sec^2\phi) \times \beta \quad (5\text{-}39)$$

In the orthogonal (perpendicular) scanning direction.

C. Scanning System One-Dimensional Relief Displacement

Like vertical aerial photographs, images acquired using an across-track scanning system contain relief displacement. However, instead of being radial from a single principal point, the displacement takes place in a direction that is perpendicular to the line of flight for each and every scan line as shown in Figure 5-17. In effect, the ground-resolution element at nadir functions like a principal point for each scan line. At nadir, the scanning system looks directly down on a tank, and it appears as a perfect circle. *The greater the height of the object above the local terrain and the greater the distance of the top of the object from nadir, the greater the amount of one-dimensional relief displacement present.* One-dimensional relief displacement is introduced in both directions away from nadir for each sweep of the across-track mirror (Jensen, 2005).

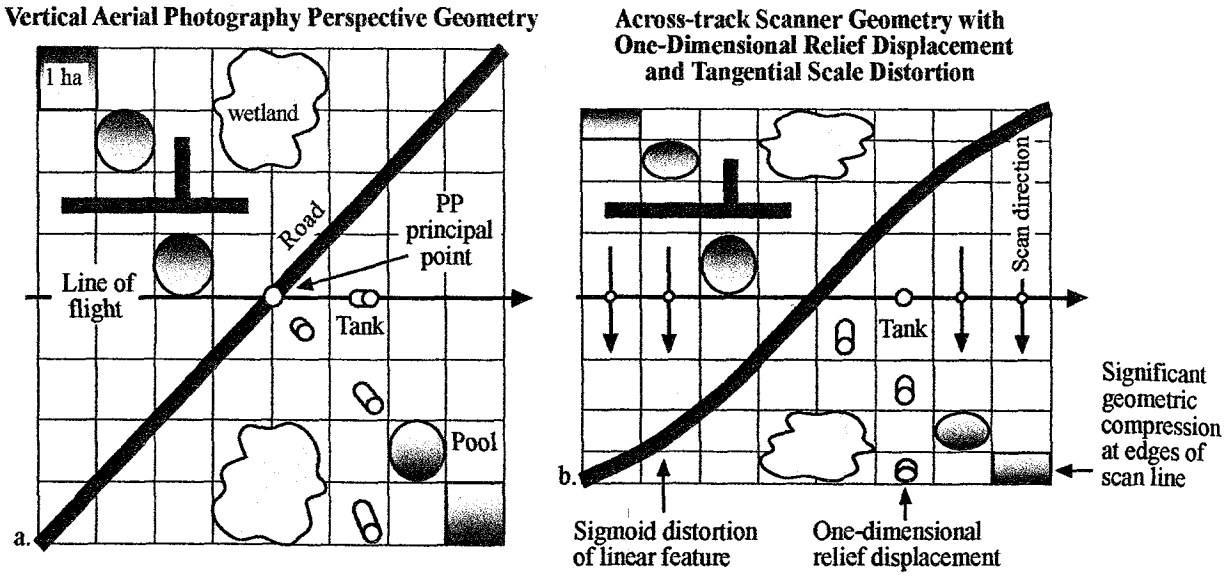


Figure 5-17: One-dimensional relief displacement and tangential scale distortion in scanning system (a) A vertical aerial photograph obtained over level terrain. Four 50-ft-tall tanks are distributed throughout the landscape and experience varying degrees of radial relief displacement the farther they are from the principal point (PP). (b) Across-track scanning system introduces one-dimensional relief displacement perpendicular to the line of flight and tangential scale distortion and compression the farther the object is from nadir. Linear features trending across the terrain are often recorded with s-shaped or sigmoid curvature characteristics due to tangential scale distortion and image compression. (From Jensen, 2005)

D. Scanning System Tangential Scale Distortion

The mirror on an across-track scanning system rotates at a constant speed and typically views from *70°* to *120°* of terrain during a complete line scan. Of course, the amount depends on the specific sensor system. The terrain directly beneath the aircraft (at nadir) is closer to the aircraft than the terrain at the edges during a single sweep of the mirror. Therefore, because the mirror rotates at a constant rate, the sensor scans a shorter geographic distance at nadir than it does at the edge of the image. This relationship tends to *compress* features along an axis that is perpendicular to the line of flight. *The greater the distance of the ground-resolution cell from nadir, the greater the image scale compression.* This is called ***tangential scale distortion.*** Objects near nadir exhibit their proper shape. Objects near the edge of the flight line become compressed and their shape distorted (Jensen, 2005). For example, consider the tangential geometric distortion and compression of the circular swimming pools and one hectare of land the farther they are from nadir in the hypothetical diagram (Figure 5-17b)

This tangential scale distortion and compression in the far range causes linear features such as roads, railroads, utility right of ways, etc., to have an *s-shape* or *sigmoid distortion* when recorded on scanner imagery. Interestingly, if the linear feature is parallel with or perpendicular to the line of flight, it does not experience sigmoid distortion (Jensen, 2005).

5.3.1.2 External Geometric Errors

External geometric errors are usually introduced by phenomena that vary in nature through space and time. The most important external variables that can cause geometric error in remote sensor data are random movements by the aircraft (or spacecraft) at the exact time of data collection, which usually involve:

- altitude changes, and/or
- attitude changes (roll, pitch, and yaw).

A. Altitude Changes

Remote sensing systems flown at a constant altitude above ground level (AGL) result in imagery with a uniform scale all along the flight line. For example, a camera with a 12-in. focal length lens flown at 20,000 ft. AGL will yield 1:20,000-scale imagery. If the aircraft or spacecraft gradually changes its altitude along a flight line, then the scale of the imagery will change (Figure 5-18a). *Increasing the altitude will result in smaller-scale imagery. Decreasing the altitude of the sensor system will result in larger-scale imagery.* The same relationship holds true for digital remote sensing systems collecting imagery on a pixel by pixel basis. The *diameter of the spot size on the ground* (*D*); the nominal spatial resolution) is a function of the instantaneous-field-of-view (β) and the altitude above ground level (*H*) of the sensor system, $D = \beta \times H$ (Jensen, 2005)

B. Attitude Changes

Satellite platforms are usually stable because they are not buffeted by atmospheric turbulence or wind. Conversely, suborbital aircraft must constantly contend with atmospheric updrafts, downdrafts, head-winds, tail-winds, and cross-winds when collecting remote sensor data. Even when the remote sensing platform maintains a constant altitude AGL, it may rotate randomly about three separate axes that are commonly referred to as *roll*, *pitch*, and *yaw* as shown in Figure 5-18b.

- Roll occurs when the aircraft or spacecraft fuselage maintains directional stability but the wings move up or down, i.e. they rotate about the x-axis angle (omega: ω). This introduces compress and/or expansion of the imagery in the near and far range perpendicular to the line of flight.

- Pitch occurs when the wings are stable but the fuselage nose or tail moves up or down, i.e., they rotate about the y-axis angle (phi: Φ). If the nose pitches down, the imagery will be compressed in the fore-direction and expanded in the aft-direction. If the nose pitches up, the imagery will be compressed in the aft-direction and expanded in the fore-direction.
- Yaw occurs when the wings remain parallel but the fuselage is forced by wind to be oriented some angle to the left or right of the intended line of flight, i.e., it rotates about the z-axis angle (kappa: κ). Thus, the plane flies straight but all remote sensor data are displaced by °κ from intended flight line (Jensen, 2005).

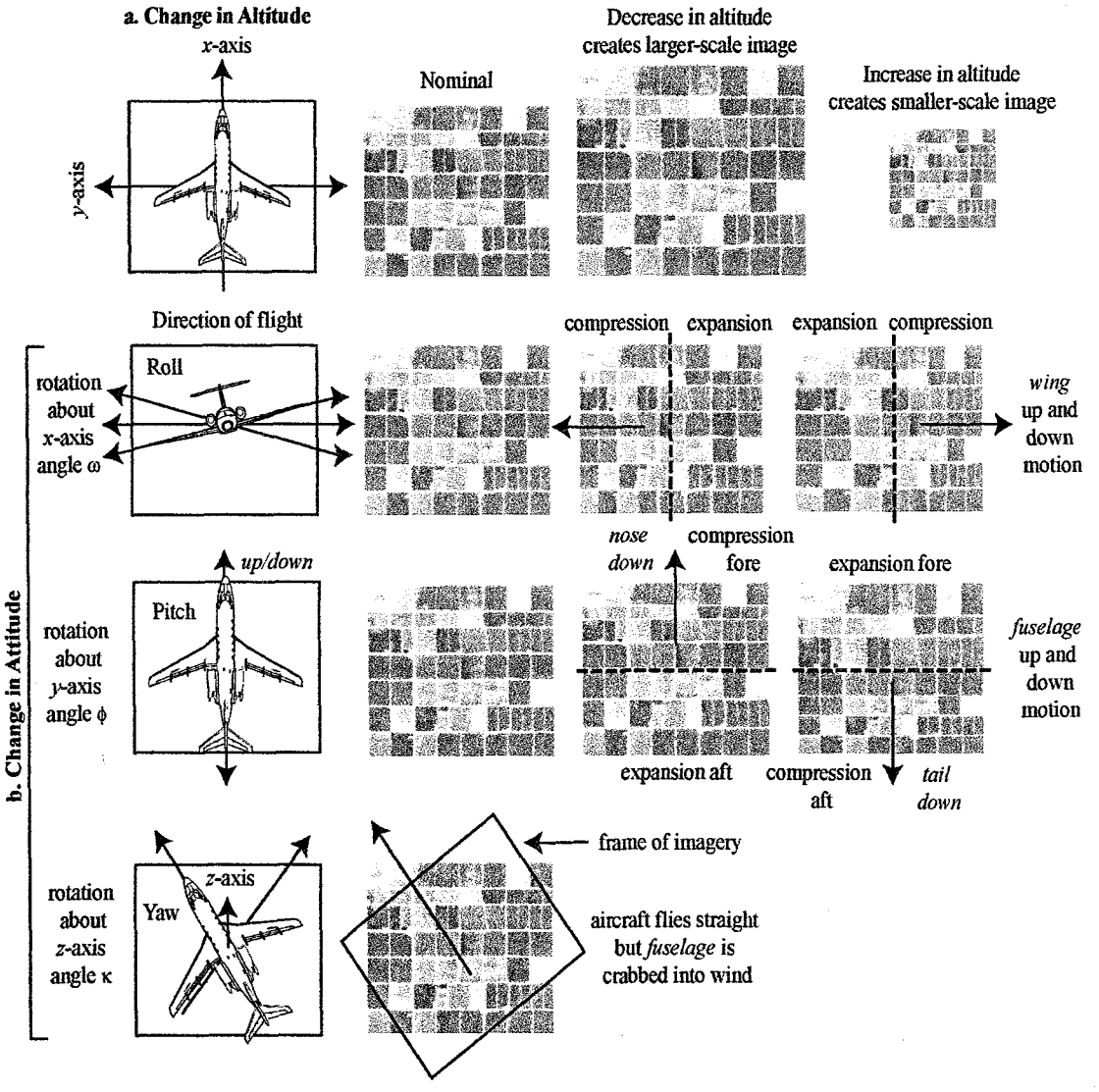


Figure 5-18: Geometric modification of remotely sensed data caused by change in platform altitude and attitude. (From Jensen, 2005).

High quality remote sensing systems have gyro-stabilization equipment that isolates the sensor system from the roll and pitch movements of the aircraft. Bernstein (1983) stated that geometric distortions introduced by sensor system *attitude* (roll, pitch, and yaw) and/or *altitude* changes can be corrected using ground control points and appropriate mathematical models. A ***ground control point*** (GCP) is a location on the surface of the Earth (e.g., a road intersection) that can be identified on the imagery and located accurately on a map. The image analyst must be able to obtain two distinct sets of coordinates associated with each GCP:

- image coordinates specified in i rows and j columns, and
- map coordinates (e.g., x,y measured in degrees of latitude and longitude, or meters in a Universal Transverse Mercator projection).

The paired coordinates (*i,j* and *x,y*) from many GCPs (e.g., 20) can be modeled to derive *geometric transformation coefficients.* These coefficients may be used to geometrically rectify the remote sensor data to a standard datum and map projection (Jensen, 2005). Several alternatives to obtaining accurate ***ground control point*** (GCP) map coordinate information for image-to-map rectification include:

- hard-copy planimetric maps;
- digital planimetric maps;
- digital orthophoto;
- Global Positioning System (GPS) instruments.

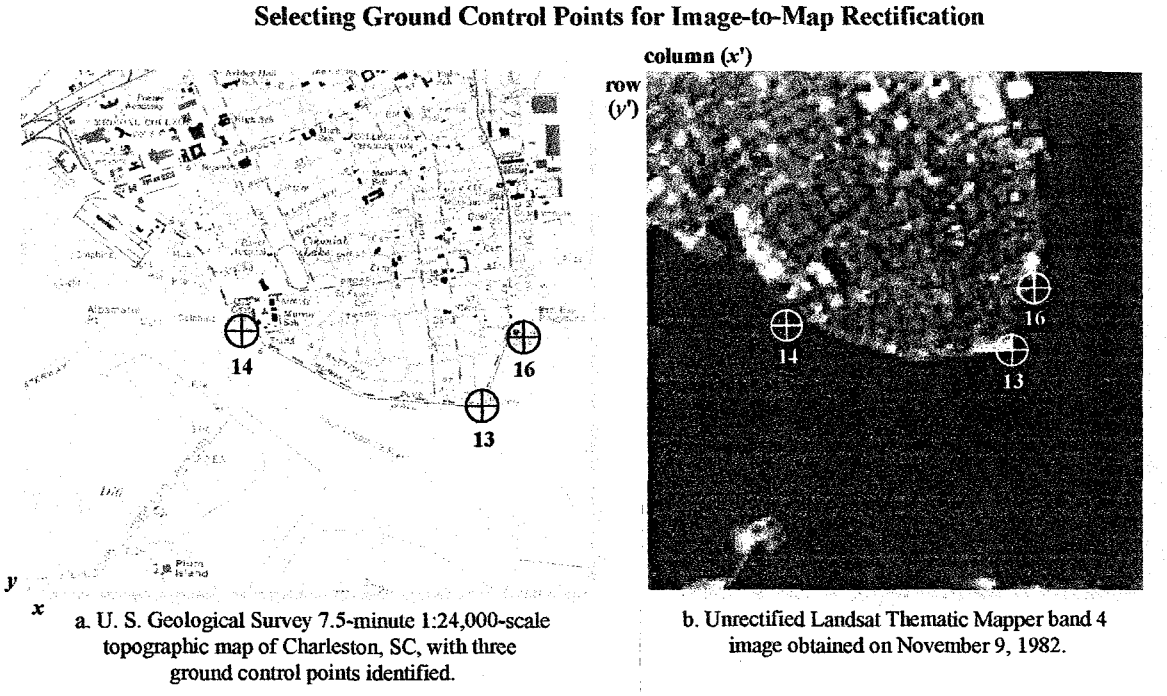
5.3.2 Types of Geometric Correction

Two common geometric correction procedures are often used by scientists to make the digital remote sensor data of value include:

- image-to-map rectification
- image-to-image registration.

5.3.2.1 Image to Map Rectification

Image-to-map rectification is the process by which the geometry of an image is made planimetric. Whenever accurate area, direction, and distance measurements are required, image-to-map geometric rectification should be performed. It may not, however, remove all the distortion caused by topographic relief displacement in images. The image-to-map rectification process normally involves selecting GCP image pixel coordinates (row and column) with their map coordinate counterparts (e.g., meters northing and easting in a Universal Transverse Mercator map projection) as shown in Figure 5-19.



a. U. S. Geological Survey 7.5-minute 1:24,000-scale topographic map of Charleston, SC, with three ground control points identified.

b. Unrectified Landsat Thematic Mapper band 4 image obtained on November 9, 1982.

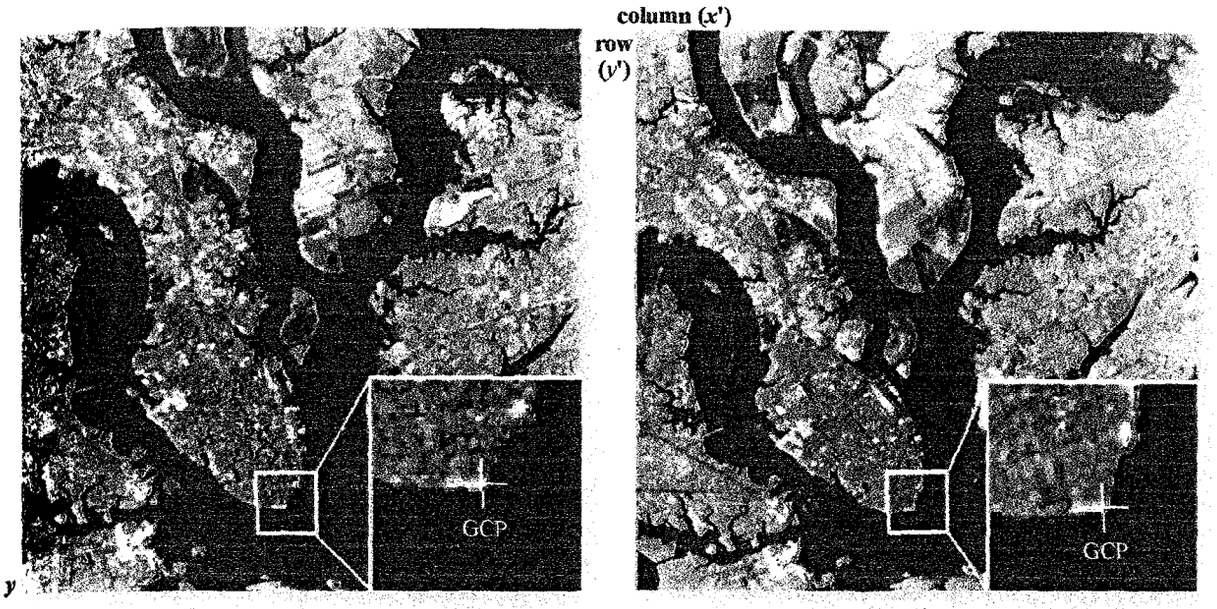
Figure 5-19: Example of image-to-map rectification. (From Jensen, 2005)

5.3.2.2 Image to Image Registration

Image-to-image registration is the translation and rotation alignment process by which two images of like geometry and of the same geographic area are positioned coincident with respect to one another so that corresponding elements of the same ground area appear in the same place on the registered images. This type of geometric correction is used when it is not necessary to have each pixel assigned a unique x, y coordinate in a map projection. For example, we might want to make a cursory examination of two images obtained on different dates to see if any change has taken place.

5.3.2.3 Hybrid Approach to Image Rectification/Registration

The same general image processing principles are used in both image rectification and image registration. If a rectified image is used as the reference base (rather than a traditional map) any image registered to it will inherit the geometric errors existing in the reference image. Because of this characteristic, most serious Earth science remote sensing research is based on analysis of data that have been rectified to a map base. However, when conducting rigorous change detection between two or more dates of remotely sensed data, it may be useful to select a hybrid approach involving both image-to-map rectification and image-to-image registration (Jensen, 2005) as shown in Figure 5-20.



a. *Rectified* Landsat TM band 4 image of Charleston, SC, obtained on November 9, 1982.

b. Unrectified Landsat TM band 4 image obtained on October 14, 1987.

Figure 5-20: Example of image-to-image hybrid registration.

5.3.3 Image to Map Geometric Rectification Logic

Two basic operations must be performed to geometrically rectify a remotely sensed image to a map coordinate system:

- Spatial interpolation,
- Intensity interpolation.

5.3.3.1 Spatial Interpolation

The geometric relationship between the input pixel coordinates (column and row; referred to as x′, y′) and the associated map coordinates of this same point (x, y) must be identified. A number of GCP pairs are used to establish the nature of the geometric coordinate transformation that must be applied to rectify or fill every pixel in the output image (x, y) with a value from a pixel in the unrectified input image (x′,y′). This process is called spatial interpolation.

Image-to-map rectification requires that polynomial equations be fit to the GCP data using least-squares criteria to model the corrections directly in the image domain without explicitly identifying the source of the distortion. Depending on the distortion in the imagery, the number of GCPs used, and the degree of topographic relief displacement in the area, higher-order polynomial equations may be required to geometrically correct the data. The order of the rectification is simply the highest exponent used in the polynomial. Figure 5-21 shows how different-order transformations fit a hypothetical surface.

This type of transformation can model six kinds of distortion in the remote sensor data (Jensen, 2005), including:

- translation in x and y,
- scale changes in x and y,
- skew, and
- rotation.

In addition, two type of filling a rectified output matrix with value from an unrectified input image includes (a) Input-to-Output (Forward) Mapping Logic and (b) Output-to-Input (Inverse) Mapping Logic as shown in Figure 5-22.

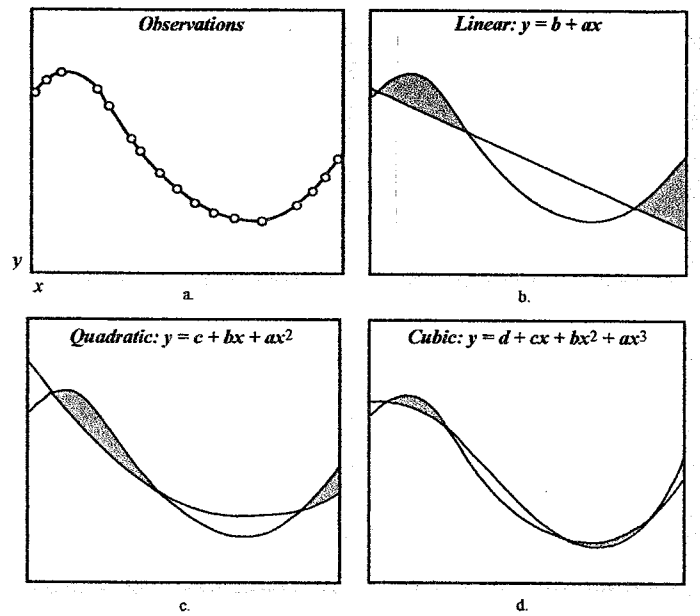


Figure 5-21: Concept of how different-order transformations fit a hypothetical surface illustrated in cross-section. a) Original observations. b) First-order linear transformation fits a plane to the data. c) Second-order quadratic fit. d) Third-order cubic fit. (From Jensen 2005).

A. Input-to-Output (Forward) Mapping

When all six operations are combined into a single expression it becomes:

$$x = a_0 + a_1 x' + a_2 y'$$

$$y = b_0 + b_1 x' + b_2 y'$$

where x and y are positions in the *output*-rectified image or map, and x' and y' represent corresponding positions in the original *input* image. These two equations can be used to perform

what is commonly referred to as *input-to-output,* or ***forward-mapping.*** As shown in Figure 5-22a, each pixel in the *input* grid (e.g., value 15 at x', y' = 2, 3) is sent to an x, y location in the output image according to the six coefficients shown.

B. Output-to-Input (Inverse) Mapping Logic

Output-to-input, or inverse mapping logic, is based on the following two equations:

$$x' = a_0 + a_1 x + a_2 y$$

$$y' = b_0 + b_1 x + b_2 y$$

where x and y are positions in the output-rectified image or map, and x′ and y′ represent corresponding positions in the original input image. The rectified output matrix consisting of x (column) and y (row) coordinates is filled in a systematic manner as shown in Figure 5-22b.Output-to-input inverse mapping logic is the preferred methodology because it results in a rectified output matrix with values at every pixel location.

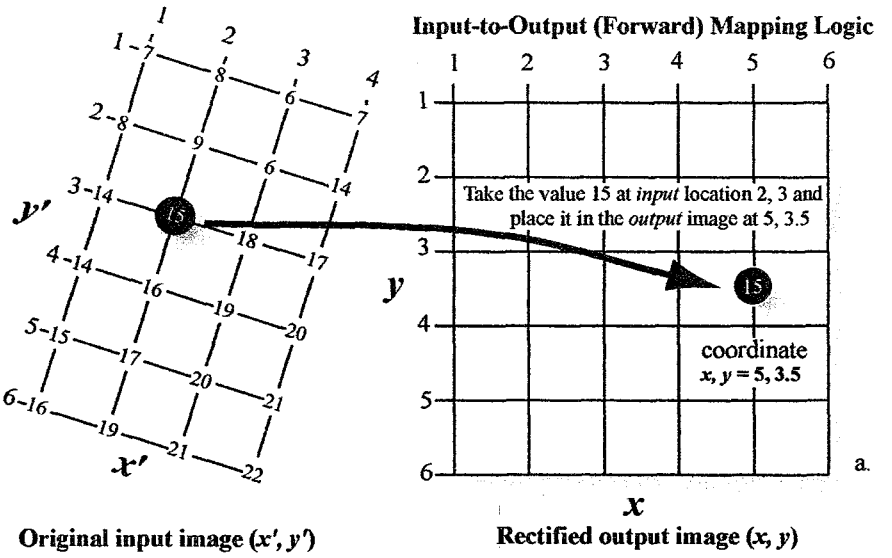
It is possible to use higher-order polynomial transformations to rectify remotely sensed data. For example, instead of using the six-parameters affine transformation, we could use a second-order (quadratic) polynomial:

$$x' = c_0 + c_1 x + c_2 y + c_3 xy + c_4 x^2 + c_5 y^2$$

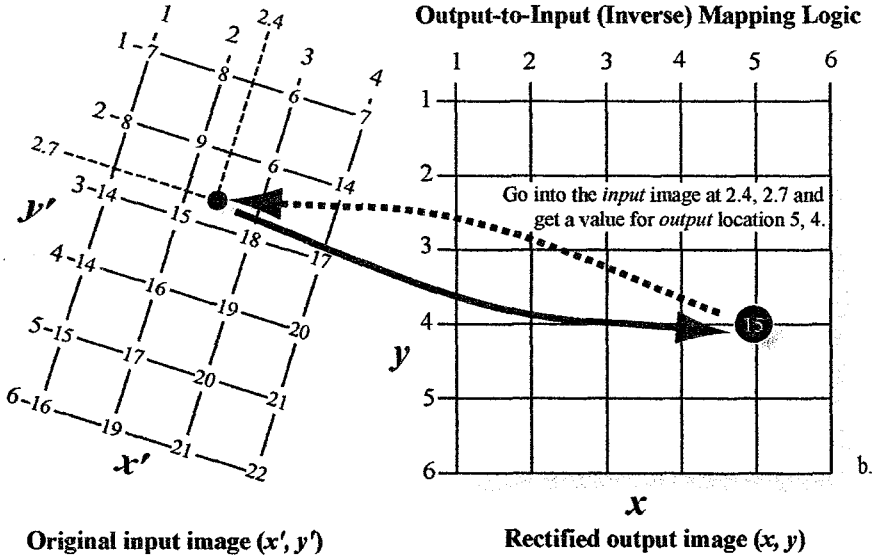
$$y' = d_0 + d_1 x + d_2 y + d_3 xy + d_4 x^2 + d_5 y^2$$

In theory, the higher the order of the polynomial, the more closely the coefficients should model the geometric error in the original input image (Figure 5-21) and place the pixels in their correct planimetric positions in the rectified output matrix. Higher-order polynomials often produce a more accurate fit for areas immediately surrounding ground control points. However, other geometric errors may be introduced at large distances from the GCPs. Figure 5-23 shows how the data are changed in linear transformations while Figure 5-24 illustrated how the data are changed in non-linear transformations.

A general rule of thumb is to use a first-order affine polynomial whenever possible. Select a higher-order polynomial (e.g., second or third order) only when there are serious geometric errors in the dataset. These types of error are often found in imagery obtained from suborbital aerial platforms where roll, pitch, and yaw by the aircraft introduce unsystematic, nonlinear distortions (Jensen, 2005).



$x = a_0 + a_1 x' + a_2 y'$
$y = b_0 + b_1 x' + b_2 y'$ where
x' and y' are locations in the original *input* image, and
x and y are locations in the rectified *output* image.



$x' = a_0 + a_1 x + a_2 y$
$y' = b_0 + b_1 x + b_2 y$

Figure 5-22: (a) The logic of filling a rectified output matrix with values from an unrectified input image matrix using *input-to-output* (***forward***) mapping logic. (b) The logic of filling a rectified output matrix with values from an unrectified input image matrix using *output-to-input* (***inverse***) mapping logic and nearest-neighbor resampling. (Form Jensen, 2005)

A simple way to measure the accuracy of a geometric rectification algorithm is to compute the Root Mean Squared Error (RMS error) for each ground control point using the equation:

$$RMS_{error} = \sqrt{(x' - x_{orig})^2 + (y' - y_{orig})^2} \quad (5\text{-}40)$$

where: xorig and yorig are the original row and column coordinates of the GCP in the image and x′ and y′ are the computed or estimated coordinates in the original image. The square root of the squared deviations represents a measure of the accuracy of each GCP in the image. By computing RMS error for all GCPs, it is possible to (1) see which GCPs contribute the greatest error, and 2) sum all the RMS error.

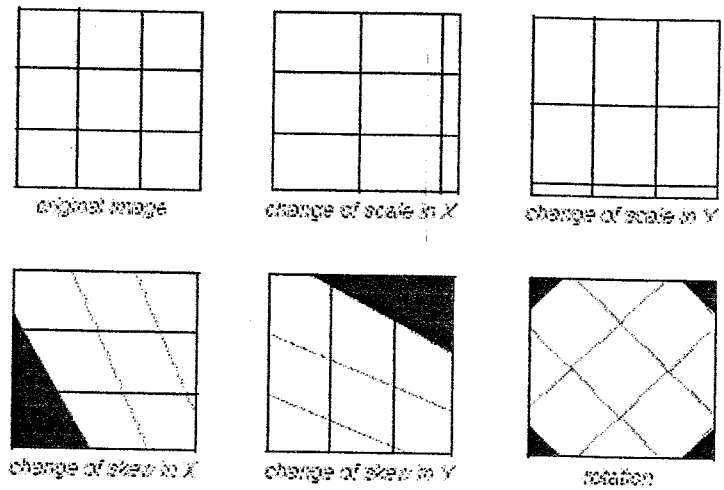


Figure 5-23: Linear transformation (From ERDAS, 2002)

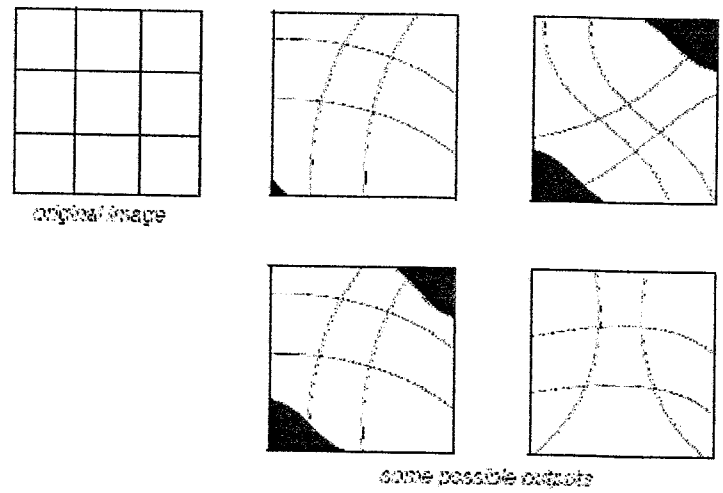


Figure 5-24: Non-linear transformation (From ERDAS, 2002)

5.3.3.2 Intensity Interpolation

Pixel brightness values must be determined. Unfortunately, there is no direct one-to-one relationship between the movement of input pixel values to output pixel locations. It will be shown that a pixel in the rectified output image often requires a value from the input pixel grid that does not fall neatly on a row-and-column coordinate. When this occurs, there must be some mechanism for determining the brightness value (BV) to be assigned to the output rectified pixel. This process is called **intensity interpolation**.

Intensity interpolation involves the extraction of a brightness value from an x′,y′ location in the original (distorted) input image and its relocation to the appropriate x, y coordinate location in the rectified output image. This pixel-filling logic is used to produce the output image line by line, column by column. Most of the time the x′ and y′ coordinates to be sampled in the input image are floating point numbers. There are several methods of brightness value (BV) interpolation include (a) nearest neighbor, (b) bilinear interpolation, and (c) cubic convolution.

A. Nearest neighbor interpolation. The brightness value closest to the predicted *x′, y′* coordinate is assigned to the output *x, y* coordinate using the *Pythagorean theorem*. As shown in Figure 5-25a, a nearest-neighbor rule would assign the output pixel (x,y) the value of *15*, which is the value found at the nearest input pixel.

B. Bilinear interpolation. Assigns output pixel values by interpolating brightness values in two orthogonal direction in the input image. It basically fits a plane to the *4* pixel values nearest to the desired position (*x′, y′*) (Figure 5-25b) and then computes a new brightness value based on the weighted distances to these points as shown in Table 5-8. The closer a pixel is to the desired x′, y′ location, the more weight it will have in the final computation of the average. The weight average of the new brightness value (BVwt) is computed using the equation.

$$Bilinear\ Interpolation\ BV_{wt} = \frac{\sum_{k=1}^{4} \frac{Z_k}{D_k^2}}{\sum_{k=1}^{4} \frac{1}{D_k^2}} \qquad (5\text{-}41)$$

where *Zk* are the surrounding four data point values, and *Dk2* are the distances squared from the point in question (***x′, y′***) to the these data points.

Table 5-8: Bilinear interpolation of a weighted brightness value at location x′, y′ based on the analysis of four sample points in Figure 5-25b

Sample Point Location (column, row)	Value at Sample Point, Z	Distance from x′, y′ to the Sample Point, D	D_k^2	$\frac{Z}{D_k^2}$	$\frac{1}{D_k^2}$
2, 2	9	$D = \sqrt{(2.4-2)^2+(2.7-2)^2} = 0.806$	0.65	13.85	1.539
3, 2	6	$D = \sqrt{(2.4-3)^2+(2.7-2)^2} = 0.921$	0.85	7.06	1.176
2, 3	15	$D = \sqrt{(2.4-2)^2+(2.7-3)^2} = 0.500$	0.25	60.00	4.000
3, 3	18	$D = \sqrt{(2.4-3)^2+(2.7-3)^2} = 0.670$	0.45	40.00	2.222
				Σ 120.91	Σ 8.937
				BV_{wt} = 120.91/8.937 =13.53	

From Jensen, 2005.

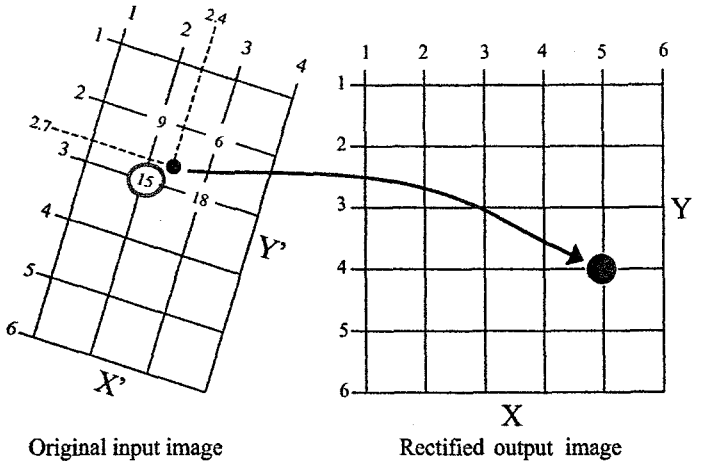
C. Cubic convolution. Assigns values to output pixels in much the same manner as bilinear interpolation, except that the weighted values of **16** pixels surrounding the location of the desired **x′, y′** pixel are used to determine the value of the output pixel. The weight average of the new brightness value (BV_{wt}) is computed using the equation:

$$Cubic\ Convolution\ BV_{wt} = \frac{\sum_{k=1}^{16} \frac{Z_k}{D_k^2}}{\sum_{k=1}^{16} \frac{1}{D_k^2}} \quad (5\text{-}42)$$

where $\mathbf{Z_k}$ are the surrounding four data point values, and $\mathbf{(D_k)^2}$ are the distances squared from the point in question (**x′, y′**) to the these data points. An example, the distances form the request x′, y′ position at 2.4, 2.7 in the input image Figure 5-25c to the closest 16 input pixel coordinates are computed in Table 5-9.

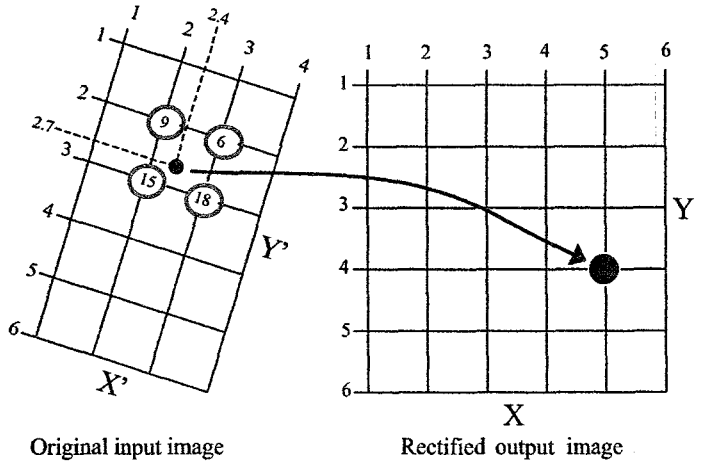
In conclusion, the image-to-map rectification process generally involved following steps:

1. selecting an appropriate planimetric base map;
2. collecting ground control points (GCPs)
3. determining the optimum set of geometric rectification coefficient by iteratively computing the total GCP RMS error, and
4. Filling the output matrix using spatial and intensity interpolation resampling methods.



Original input image Rectified output image

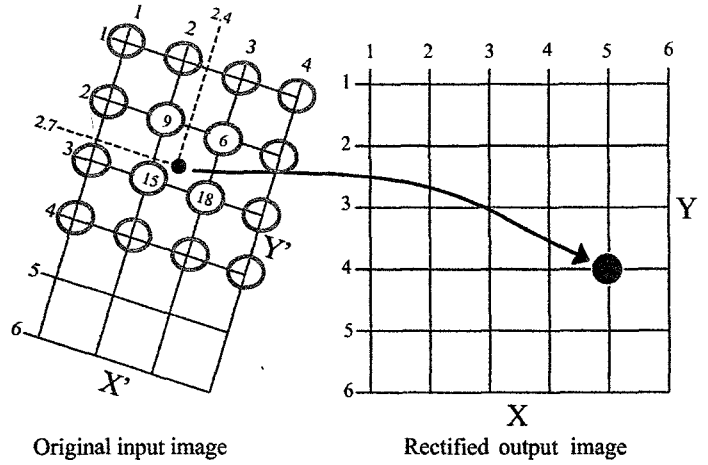
(a) Nearest neighbor Interpolation algorithm



$$BV_{wt} = \frac{\sum_{k=1}^{4} \frac{Z_k}{D_k^2}}{\sum_{k=1}^{4} \frac{1}{D_k^2}}$$

Original input image Rectified output image

(b) Bilinear Interpolation algorithm



$$BV_{wt} = \frac{\sum_{k=1}^{16} \frac{Z_k}{D_k^2}}{\sum_{k=1}^{16} \frac{1}{D_k^2}}$$

Original input image Rectified output image

(c) Cubic Convolution algorithm

Figure 5-25: Spatial interpolation algorithm (Modified from Jensen, 2005).

Table 5-9: Cubic convolution interpolation of a weighted brightness value at location x′, y′ based on the analysis of 16 sample points in Figure 5-26c

Sample Point Location (column, row)	Value at Sample Point, Z	Distance from x', y' to the Sample Point, D	D_k^2	$\frac{Z}{D_k^2}$	$\frac{1}{D_k^2}$
1, 1	7	$D = \sqrt{(2.4-1)^2+(2.7-1)^2} = 2.202$	4.85	1.443	0.206
2, 1	8	$D = \sqrt{(2.4-2)^2+(2.7-1)^2} = 1.746$	3.05	2.623	0.328
3, 1	6	$D = \sqrt{(2.4-3)^2+(2.7-1)^2} = 1.80$	3.24	1.852	0.309
4, 1	7	$D = \sqrt{(2.4-4)^2+(2.7-1)^2} = 2.335$	5.45	1.284	0.183
1, 2	8	$D = \sqrt{(2.4-1)^2+(2.7-2)^2} = 1.565$	2.45	3.265	0.408
2, 2	9	$D = \sqrt{(2.4-2)^2+(2.7-2)^2} = 0.806$	0.65	13.85	1.539
3, 2	6	$D = \sqrt{(2.4-3)^2+(2.7-2)^2} = 0.921$	0.85	7.06	1.176
4, 2	14	$D = \sqrt{(2.4-4)^2+(2.7-2)^2} = 1.746$	3.05	4.59	0.328
1, 3	14	$D = \sqrt{(2.4-1)^2+(2.7-3)^2} = 1.432$	2.05	6.829	0.488
2, 3	15	$D = \sqrt{(2.4-2)^2+(2.7-3)^2} = 0.500$	0.25	60.00	4.000
3, 3	18	$D = \sqrt{(2.4-3)^2+(2.7-3)^2} = 0.670$	0.45	40.00	2.222
4, 3	17	$D = \sqrt{(2.4-4)^2+(2.7-3)^2} = 1.63$	2.65	6.415	0.377
1, 4	14	$D = \sqrt{(2.4-1)^2+(2.7-4)^2} = 1.911$	3.65	3.836	0.274
2, 4	16	$D = \sqrt{(2.4-2)^2+(2.7-4)^2} = 1.360$	1.85	8.649	0.541
3, 4	19	$D = \sqrt{(2.4-3)^2+(2.7-4)^2} = 1.432$	2.05	9.268	0.488
4, 4	20	$D = \sqrt{(2.4-4)^2+(2.7-4)^2} = 2.062$	4.25	4.706	0.235
				Σ 175.67	Σ 13.102

$BV_{wt} = 175.67 / 13.102$

$BV_{wt} = \mathbf{13.41}$

From Jensen, 2005.

Chapter 6: Image Enhancement in Digital Image Processing

Image enhancement is the process of making an image more interpretable for a particle application (Faust, 1989). Image enhancement algorithms are applied to remotely sensed data to improve the appearance of an image for human visual analysis or occasionally for subsequent machine analysis. There is no such thing as the ideal or best image enhancement because the results are ultimately evaluation by humans, who make subjective judgments as to whether a given image enhancement is useful (Jensen, 2005).

The techniques to be used in image enhancement depend upon (ERDAS, 2002:

- Your data—the different bands of Landsat, SPOT, and other imaging sensors are selected to detect certain features. You must know the parameters of the bands being used before performing any enhancement.
- Your objective—for example, sharpening an image to identify features that can be used for training samples requires a different set of enhancement techniques than reducing the number of bands in the study. You must have a clear idea of the final product desired before enhancement is performed.
- Your expectations—what you think you are going to find.
- Your background—your experience in performing enhancement.

Basically, image enhancement techniques can be divided into two basic categories: point and neighborhood. ***Point techniques*** enhance the pixel based only on its value, with no concern for the values of neighboring pixels. These techniques include contrast stretches (non-adaptive), classification, and level slices. ***Neighborhood techniques*** enhance a pixel based on the values of surrounding pixels. As a result, these techniques require the processing of a possibly large number of pixels for each output pixel. The most common way of implementing these enhancements is via a moving window convolution.

ERDAS (2002) classified image enhancement into three categories included (1) radiometric enhancement, (2), spatial enhancement and (3) spectral enhancement. In addition, McCloy (2006) added one more category, "temporary enhancement", in enhancement operation. The major techniques for each category can be summarized based on McCloy (2006) as follow:

(1). Radiometric Enhancement. The recorded radiometric data are enhanced to provide better grey tone or color condition in the viewed image.

- Display of an image
- Pseudo-color Density Slice
- Linear enhancement

- Non-linear enhancement
- Piecewise Linear Stretch

(2). Spatial Enhancement. To exact information on spatial inter-relationships or the texture and pattern in the data.

- Measurement of Texture
- Edge Detection
- Removal of Regular Noise in Image Data
- Analysis of Spatial Correlation: The Variogram
- Image Segmentation
- Object Patterns and Object Sizes: The ALV Function

(3). Spectral enhancement. The spectral channels are combined or transformed so as to enhance certain features in the data.

- Ratioing
- Orthogonal Transformations
- Vegetation indices
- Fourier Transformation

(4). Temporal enhancement. To compress a temporal sequence of images, yet to extract information on the dynamic changes that are recorded in the image data.

- Temporal Enhancement
- Principal Component
- Temporal Distance Image
- Fourier Analysis

In this chapter, selected image enhancement operations that have proven of value for visual analysis and/or subsequent digital image processing will be here reviewed in three categories including (1) radiometric enhancement, (2), spatial enhancement and (3) spectral enhancement.

6.1 Radiometric Enhancement

Basically, radiometric enhancement deals with the individual values of the pixels in the image. Radiometric enhancement that are applied to one band may not be appropriate for other bands depending on the point and bands in which appear. Therefore, the radiometric enhancement of a multi-band image can usually be considered as a series of independent, single band enhancements (Faust, 1989). Radiometric enhancement usually does not bring out the contrast of every pixel in an image. Contrast can be lost between some pixels, while gain on others (ERDAS, 2002) as shown in Figure 6-1.

In Figure 6-1, the range between j and k in the histogram of the original data is about one third of the total range of the data. When the same data are radiometrically enhanced, the range between j and k can be widened. Therefore, the pixels between j and k gain contrast—it is easier to distinguish different brightness values in these pixels. Conversely, the pixels outside the range between j and k are more grouped together than in the original histogram to compensate for the stretch between j and k. Contrast among these pixels is lost.

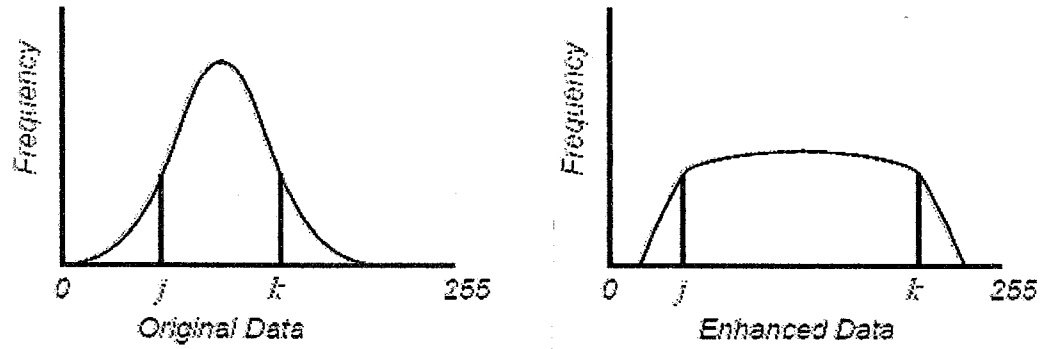


Figure 6-1: Histograms of radiometrically enhanced data (From, ERDAS, 2002).

When radiometric enhancements are performed on the display device, the transformation of data file values into brightness values is illustrated by the graph of a lookup table as shown in Figure 6-2. It shows the graph of a lookup table that increases the contrast of data file values in the middle range of the input data (the range within the brackets). Note that the input range within the bracket is narrow, but the output brightness values for the same pixels are stretched over a wider range. This process is called **contrast stretching** (ERDAS, 2002).

Basically, there are linear and nonlinear digital contrast enhancements. The terms linear and nonlinear refer to the function that is applied to the data to perform the enhancement as shown in Figure 6-3. A piecewise linear stretch uses a polyline function to increase contrast to varying degrees over different ranges of the data.

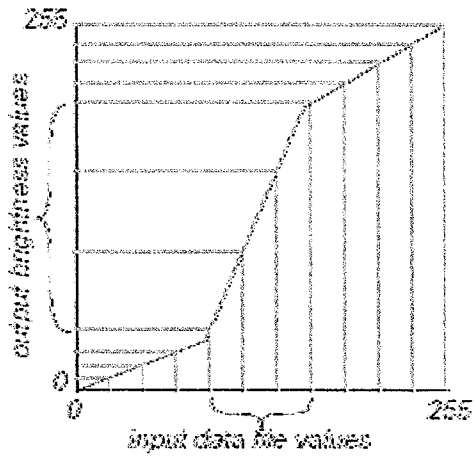


Figure 6-2: Graph of a Lookup Table (From ERDAS, 2002)

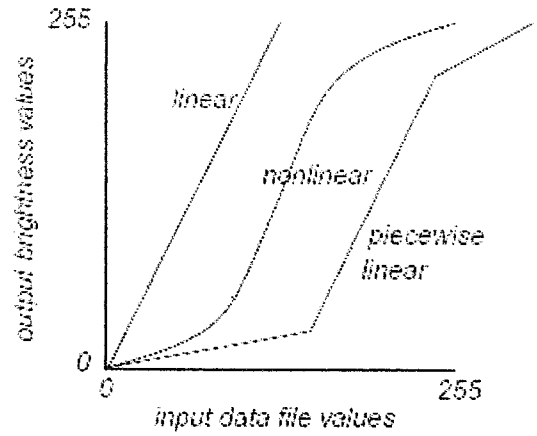


Figure 6-3: Enhancement with Lookup Tables (From ERDAS, 2002)

6.1.1. Linear Contrast Enhancement

A linear contrast stretch is a simple way to improve the visible contrast of an image. It is often necessary to contrast-stretch raw image data, so that they can be seen on the display. In most raw data, the data file values fall within a narrow range—usually a range much narrower than the display device is capable of displaying. That range can be expanded to utilize the total range of the display device (usually 0 to 255).

Jensen (2005) summarized commonly use of linear contrast enhancement techniques including

- Minimum-Maximum Contrast Stretching;
- Percentage Linear and Standard Deviation Stretching;
- Piecewise Linear Contrast Stretching.

A. Minimum- Maximum Contrast Stretching

To perform a linear contrast enhancement, analysts examine the image statistics and determine the minimum and maximum brightness values in band k, min_k and max_k, respectively. The output brightness value, BV_{out}, is computed as:

$$BV_{out} = \left(\frac{BV_{in} - \min_k}{\max_k - \min_k} \right) quant_k \qquad (6\text{-}1)$$

where: BVin s the original input brightness value

$quant_k$ s the range of the brightness values (e.g. 8 bits equals 255),

min_k s the minimum value in the image,

max_k s the maximum value in the image, and

BV_{out} s the output brightness value

Thus, the original brightness values between 4 and 104 are linearly distributed between 0 and 255 (Figure 6-4) by computing based on Equation 6-1 as following:

$$BV_{out} = \left(\frac{4_{in} - 4_{\min}}{104_{\max} - 4_{\min}} \right) 255 = 0$$

$$BV_{out} = \left(\frac{104_{in} - 4_{\min}}{104_{\max} - 4_{\min}} \right) 255 = 255$$

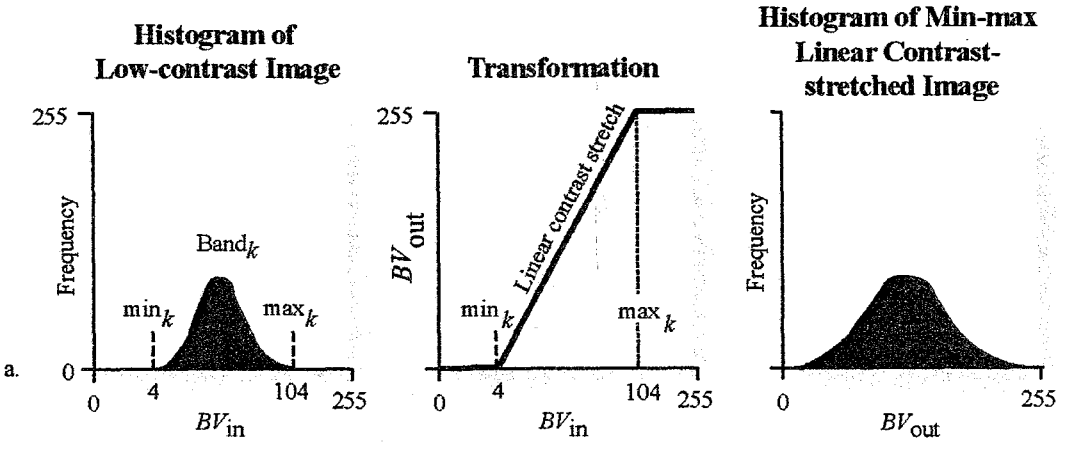


Figure 6-4: Minimum-Maximum Contrast Stretching (From Jensen, 2005)

B. Percentage Linear and Standard Deviation Contrast Stretching

Image analysts often specify min_k and max_k that lie a certain percentage of pixel from the mean of the histogram. This is called a *percentage linear contrast stretch.* If the percentage coincides with a standard deviation percentage, then it is called a *standard deviation contrast stretch.* For normal distribution, 68 % of the observations lie within ± 1 standard deviation of the mean, 95.4 % of all observations lie within ± 2 standard deviations, and 99.73 % within ± 3 standard deviations. Figure 6-5 shows the results of applying a ± standard deviation percentage linear contrast stretch. This moves the min_k and max_k values ± 34% from the mean into the tails of the distribution.

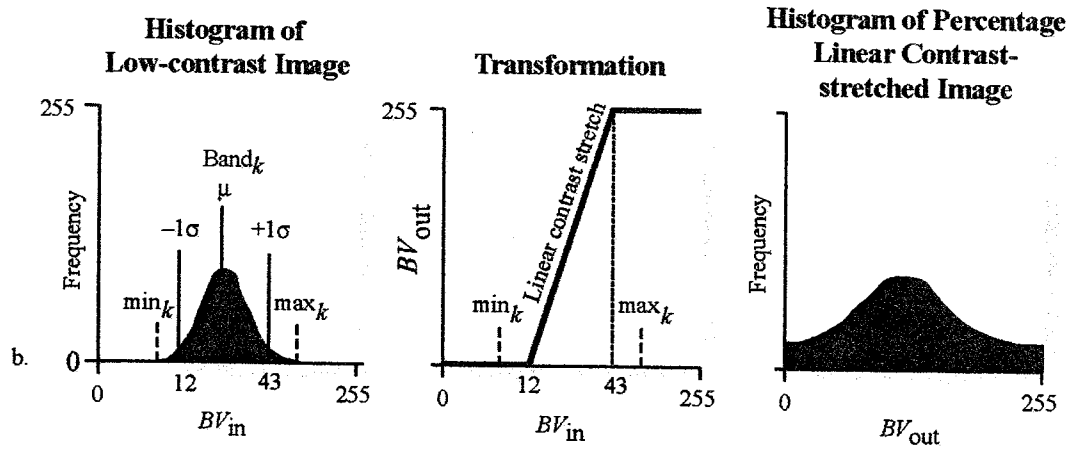


Figure 6-5: Percentage Linear and Standard Deviation Contrast Stretching (From Jensen, 2005)

C. Piecewise Linear Contrast Stretching

A piecewise linear contrast stretch allows for the enhancement of a specific portion of data by dividing the lookup table into three sections: low, middle, and high. It enables you to create a number of straight line segments that can simulate a curve. You can enhance the contrast or brightness of any section in a single color gun at a time. This technique is very useful for enhancing image areas in shadow or other areas of low contrast (ERDAS, 2002). A piecewise linear contrast stretch normally follows two rules:

- The data values are continuous; there can be no break in the values between High, Middle, and Low. Range specifications adjust in relation to any changes to maintain the data value range.
- The data values specified can go only in an upward, increasing direction (Figure 6-6).

Jensen (2005) stated that when the histogram of an image is not Gaussian, it is possibly to apply a piecewise linear contrast stretch to the imagery. In effect, this corresponds to setting up a series of min_k and max_k and using Equation 6-1 within user-selected regions of the histogram.

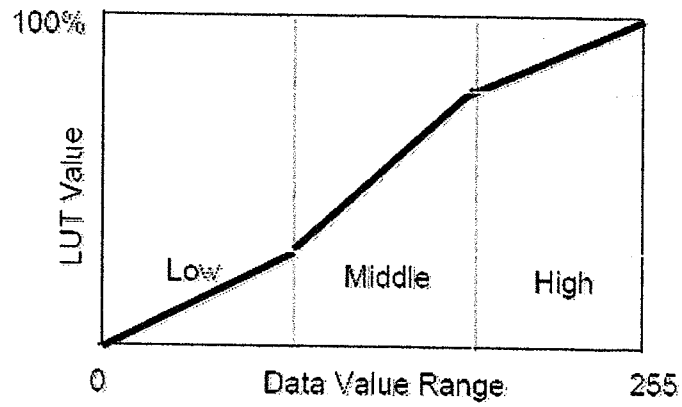


Figure 6-6: Piecewise Linear Contrast Stretch (From ERDAS, 2002)

The contrast value for each range represents the percent of the available output range that particular range occupies. The brightness value for each range represents the middle of the total range of brightness values occupied by that range. Since rules 1 and 2 above are enforced, as the contrast and brightness values are changed, they may affect the contrast and brightness of other ranges. For example, if the contrast of the low range increases, it forces the contrast of the middle to decrease (ERDAS, 2002).

6.1.2 Nonlinear Contrast Stretch.

A nonlinear spectral enhancement can be used to gradually increase or decrease contrast over a range, instead of applying the same amount of contrast (slope) across the entire image. Usually, nonlinear enhancements bring out the contrast in one range while decreasing the contrast in other ranges as shown in Figure 6-7.

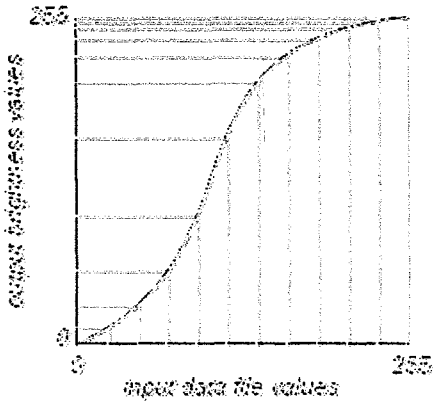


Figure 6-7: Nonlinear radiometric enhancement (From ERDAS, 2002)

A. Histogram Equalization

Jensen (2005) claimed that one of the most useful nonlinear contrast enhancements is histogram equalization. The algorithm passes through the individual bands of the dataset and assigns approximately an equal number of pixels to each of the user-specified output grey scale classes (i.e., 32, 64, 256). Histogram equalization applied the greatest contrast enhancement to the most populated range of brightness values in the image. It automatically reduces the contrast in the very light or dark parts of the image associated with the tails of a normally distributed histogram as shown in Figure 6-8. Histogram equalization can also separate pixels into distinct groups if there are few output values over a wide range. This can have the visual effect of a crude classification (ERDAS, 2002).

Schowengerdt (1997) mentioned that the contrast of an equalized image is often rather harsh, so equalization is not recommended as a general purpose stretch, However, no parameter are required from the analyst to implement the transformation making it easy to apply.

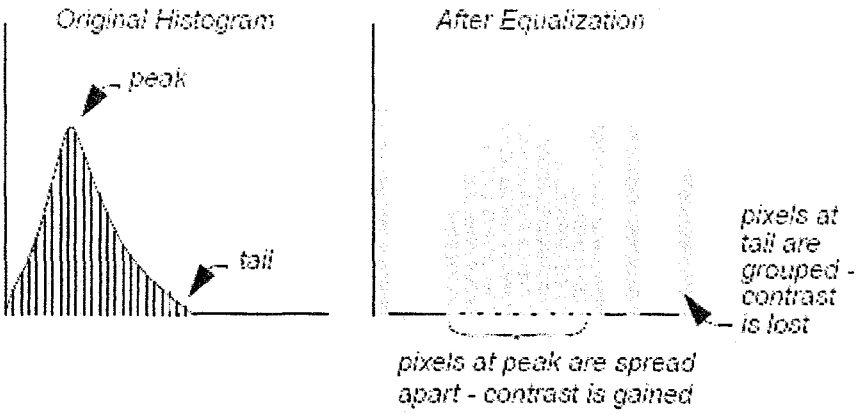


Figure 6-8: Histogram Equalization (From ERDAS, 2002).

To perform a histogram equalization, the pixel values of an image (either data file values or brightness values) are reassigned to a certain number of bins, which are simply numbered sets of pixels. The pixels are then given new values, based upon the bins to which they are assigned (ERDAS, 2002). The following parameters are entered:

- N is the number of bins to which pixel values can be assigned. If there are many bins or many pixels with the same value(s), some bins may be empty.
- M is the maximum of the range of the output values. The range of the output values is from 0 to M.

The total number of pixels is divided by the number of bins, equaling the number of pixels per bin, as shown in the following equation:

$$A = \frac{T}{N} \tag{6-2}$$

Where:

N = the number of bins

T = the total number of pixels in the image

A = the equalized number of pixels per bin

The pixels of each input value are assigned to bins, so that the number of pixels in each bin is as close to A as possible. Consider Figure 6-9, there are 240 pixels represented by histogram. To equalize this histogram to 10 bins, there would be: 240 pixels / 10 bins = 24 pixels per bin = A. To assign pixels to bins, the following equation is used:

$$B_i = \text{int}\left|\frac{\left(\sum_{k=1}^{i-1} H_k\right) + \frac{H_i}{2}}{A}\right| \quad (6\text{-}3)$$

Where:

A is an equalized number of pixels per bin (see above),

H_i is the number of values with the value i (histogram),

int is the integer function (truncating real numbers to integer),

B_i is bin number for pixels with value i.

The 10 bins are rescaled to the range 0 to M. In this example, M = 9, because the input values ranged from 0 to 9, so that the equalized histogram can be compared to the original. The output histogram of this equalized image looks like Figure 6-10. Comparison radiometric enhancement by major linear and nonlinear is shown in Figure 6-11.

B. Brightness Inversion

The brightness inversion functions produce images that have the opposite contrast of the original image. Dark detail becomes light, and light detail becomes dark. This can also be used to invert a negative image that has been scanned to produce a positive image (ERDAS, 2002).

Brightness inversion has two options: inverse and reverse. (Figure 6-12). Both options convert the input data range (commonly 0 - 255) to 0 - 1.0. A min-max remapping is used to simultaneously stretch the image and handle any input bit format. The output image is in floating point format, so a min-max stretch is used to convert the output image into 8-bit format. Inverse is useful for emphasizing detail that would otherwise be lost in the darkness of the low DN pixels. This function applies the following algorithm:

$$DN_{out} = 1.0 \text{ if } 0.0 < DN_{in} < 0.1 \quad (6\text{-}4)$$

$$DN_{out} = 0.1 \div DN_{in} \text{ if } 0.1 < DN_{in} < 1 \quad (6\text{-}5)$$

Reverse is a linear function that simply reverses the DN values:

$$DN_{out} = 1.0 - Dn_{in} \quad (6\text{-}6)$$

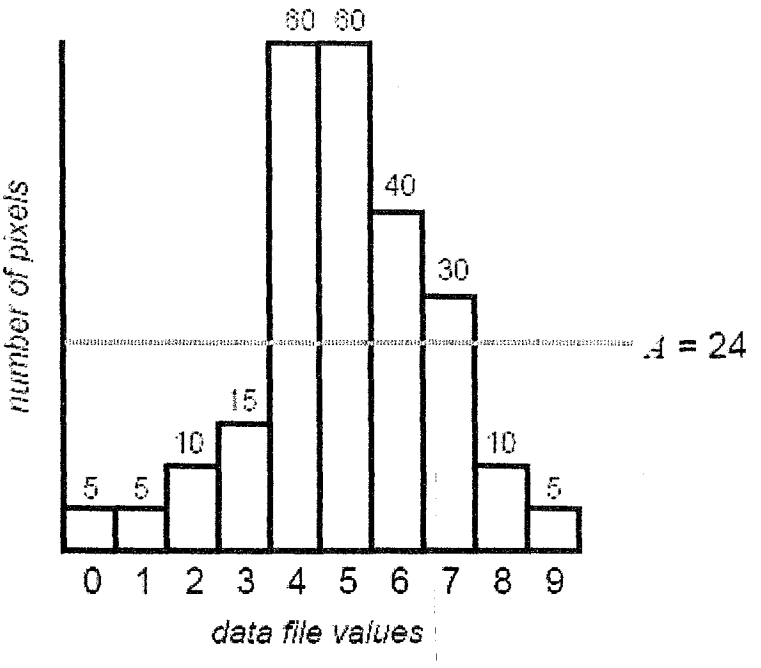


Figure 6-9: Original Histogram Data (From ERDAS, 2002)

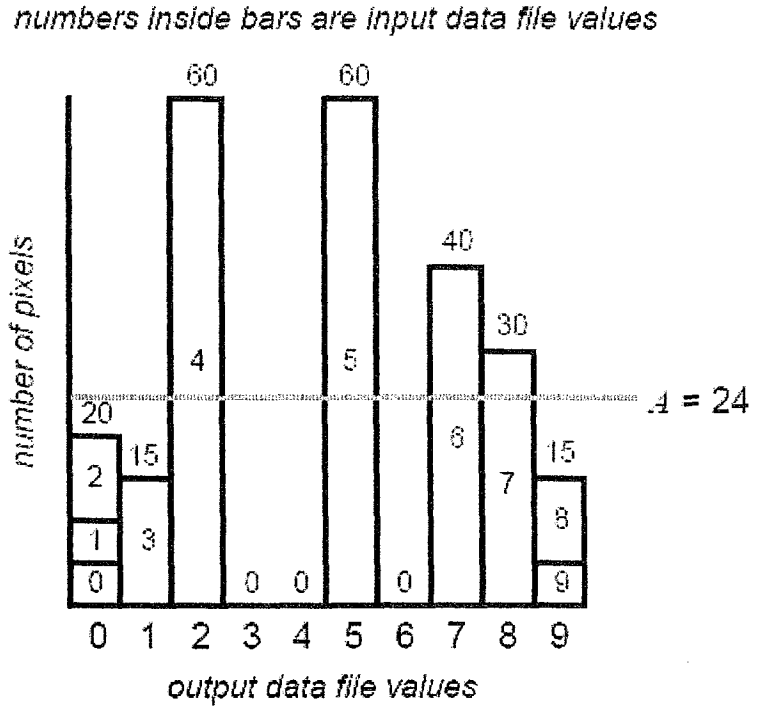
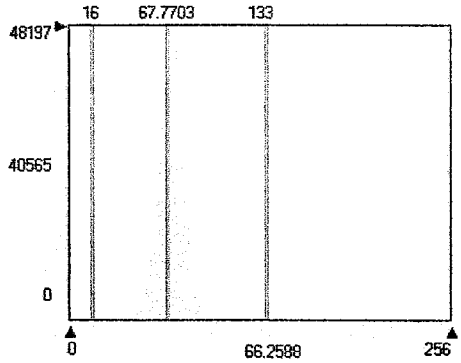
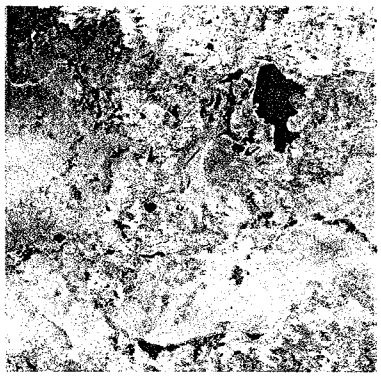
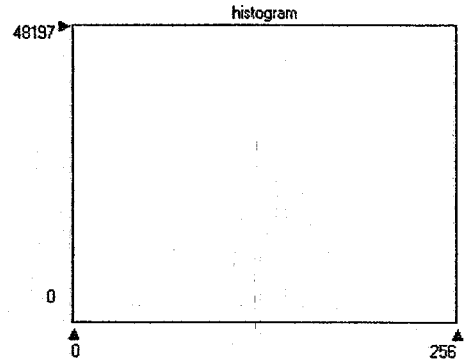


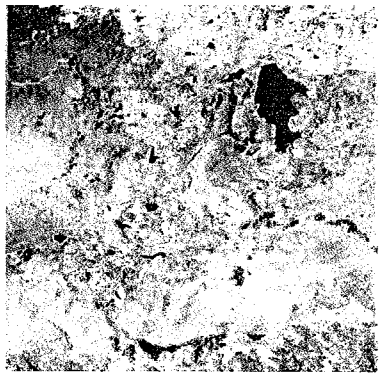
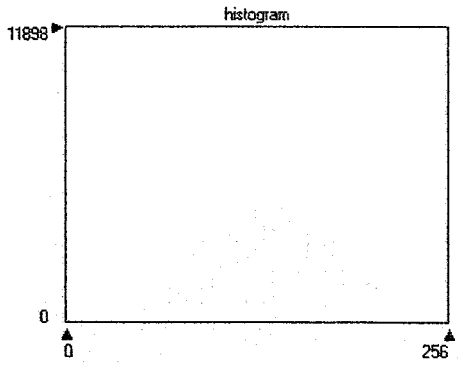
Figure 6-10: Equalized Histogram Data (From ERDAS, 2002).

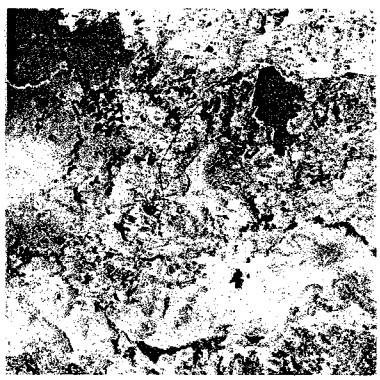
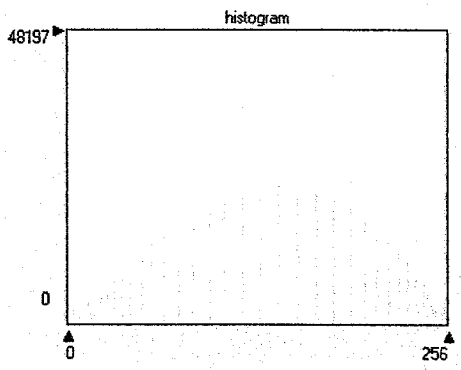
(a) Original Data Landsat TM Band 4

(b) Min. – Max. Contrast Stretching

(c) 2 Std. Deviation Contrast Stretching

(d) Histogram Equalization Contrast Stretching

Figure 6-11: Comparison of radiometric enhancement of Landsat TM band 4, Date 6 March 2005, Non Din Deang, Buriram, Thailand: (a) Original (b) Minimum-maximum contrast stretching (c) Two

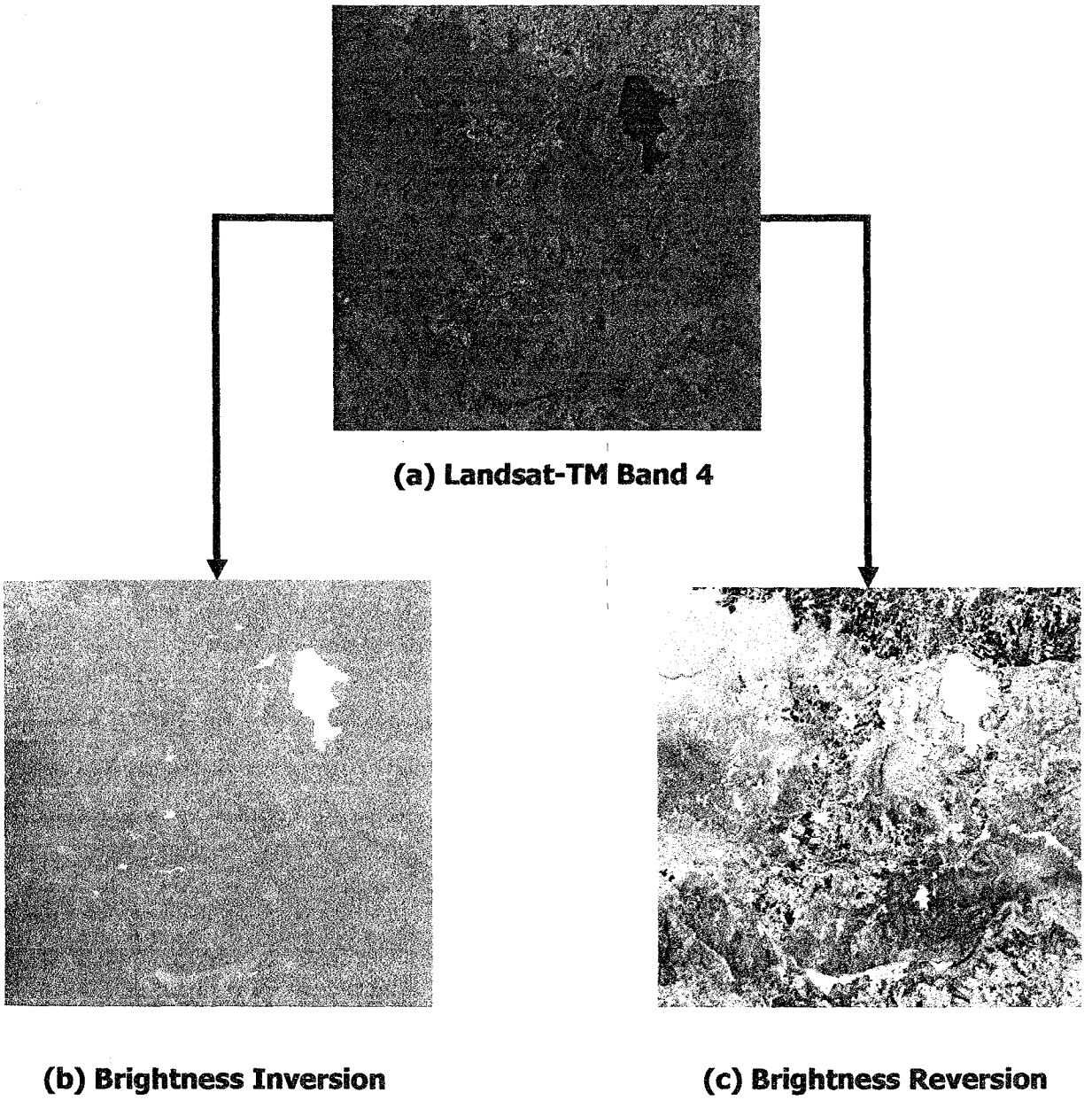


Figure 6-12: Example of Brightness Inversion of Landsat TM band 4, Date 6 March 2005, Non Din Deang, Buriram, Thailand: (a) Original brightness data (b) Brightness inversion data and (c) Brightness reversion data.

6.2 Spatial Enhancement

Spatial enhancement modified pixel values based on the values of surrounding pixels. Spatial enhancement deals largely with spatial frequency, which is difference between the highest and lowest values of a contiguous set of pixels. Jensen (2005) defines spatial frequency as "the number of changes in brightness value per unit distance for any particular part of an image" as shown in Figure 6-13.

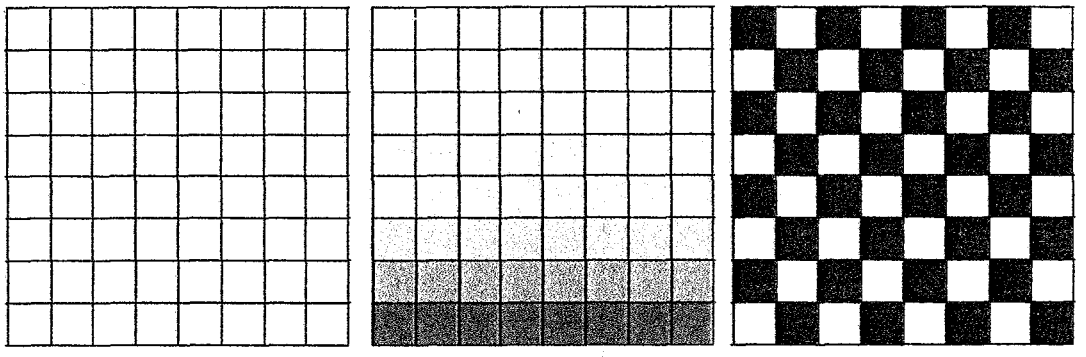


Figure 6-13: Spatial Frequencies (From ERDAS, 2002).

Based on spatial frequencies in Figure 6-13, it also implies about contrast as follows:

- zero spatial frequency (No contrast) - a flat image, in which every pixel has the same value
- low spatial frequency (Low contrast) - an image consisting of a smoothly varying gray scale
- highest spatial (High Contrast) frequency - an image consisting of a checkerboard of black and white pixels

A characteristic of remotely sensed image is a parameter called spatial frequency; defined as the number of changes in brightness value per unit distance for any particular part of an image. If there are very few changes in brightness value over a given area in an image, this is commonly referred to as a low-frequency area. Conversely, if the brightness values change dramatically over short distance, this is an area of high-frequency detail.

Spatial frequency in remotely sensed imagery may be enhanced or subdued using two different approaches (Jensen, 2005).

1. Spatial convolution filtering based primarily on the use of convolution masks. The procedure is relatively easy to understand and can be used to enhance low-and-high frequency detail, as well as edges in the imagery.
2. Fourier analysis, which mathematically separates an image into its spatial frequency component. It is possible interactively to emphasize certain groups or bands of frequencies relative to others and recombine the spatial frequencies to produce an enhanced image.

6.2.1 Spatial Convolution Filtering

Basically, a linear spatial filter is a filter for which the brightness value ($BV_{i,j}$) at location *i,j* in the output image is a function of some weighted average (linear combination) of brightness values located in a particular spatial pattern around the *i,j* location in the input image. The process of evaluating the weighted neighboring pixel values is called two-dimensional convolution filtering (Jensen, 2005). The procedure is often used to change the spatial frequency characteristics of an image. For example, a linear spatial filter that emphasizes high spatial frequencies may sharpen the edges within an image. A linear spatial filter that emphasizes low spatial frequencies may be used to noise within an image.

A convolution kernel is a matrix of numbers that is used to average the value of each pixel with the values of surrounding pixels in a particular way. The numbers in the matrix serve to weight this average toward particular pixels. (Figure 6-14) These numbers are often called coefficients, because they are used as such in the mathematical equations.

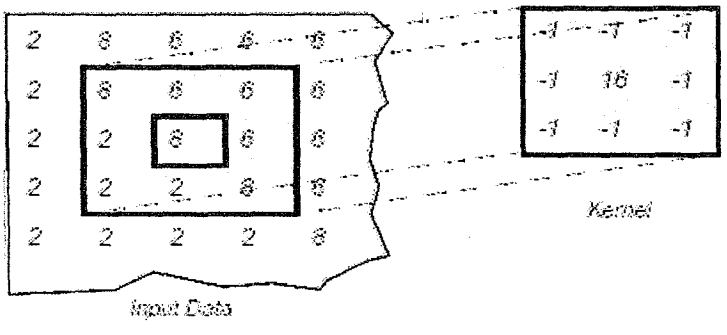


Figure 6-14: Applying a Convolution Kernel (From ERDAS, 2002).

The equation is used to derive an output data file value for the pixel being convolved (in the center) (ERDAS, 2002) as following:

$$V = \left| \frac{\sum_{i=1}^{q} \left(\sum_{j=1}^{q} f_{ij} d_{ij} \right)}{F} \right| \qquad (6\text{-}7)$$

where:

f_{ij} is the coefficient of a convolution kernel at position *i,j* (in the kernel)

d_{ij} is the data value of the pixel that corresponds to f_{ij}

q is the dimension of the kernel, assuming a square kernel (if q = 3, the kernel is 3 × 3)

F is either the sum of the coefficients of the kernel, or 1 if the sum of coefficients is 0

V is the output pixel value

In cases where V is less than 0, V is clipped to 0. The sum of the coefficients (*F*) is used as the denominator of the equation above, so that the output values are in relatively the same range as the input values. Since *F* cannot equal zero (division by zero is not defined), *F* is set to 1 if the sum is zero.

Fundamentally, three types of kernels are applied for spatial convolution filtering. The major characterization of each type was summarized by ERDAS (2002) as following:

(1) Low-Frequency Kernels. Below is an example of a low-frequency kernel, or low-pass kernel, which decreases spatial frequency:

1	1	1
1	1	1
1	1	1

This kernel simply averages the values of the pixels, causing them to be more homogeneous. The resulting image looks either more smooth or more blurred.

(2) High-Frequency Kernels. A high-frequency kernel, or high-pass kernel, has the effect of increasing spatial frequency. High-frequency kernels serve as edge enhancers, since they bring out the edges between homogeneous groups of pixels. Unlike edge detectors (such as zero-sum kernels), they highlight edges and do not necessarily eliminate other features.

-1	-1	-1
-1	16	-1
-1	-1	-1

When this kernel is used on a set of pixels in which a relatively low value is surrounded by higher values, the low value gets lower as result below:

BEFORE			AFTER		
204	200	197	204	200	197
201	106	209	201	9	209
198	200	210	198	200	210

Inversely, when the kernel is used on a set of pixels in which a relatively high value is surrounded by lower values, the high value becomes higher as result below:

BEFORE			AFTER		
64	60	57	64	60	57
61	125	69	61	187	69
58	60	70	58	60	70

In either case, spatial frequency is increased by this kernel.

(3) Zero-Sum Kernels. Zero-sum kernels are kernels in which the sum of all coefficients in the kernel equals zero. When a zero-sum kernel is used, then the sum of the coefficients is not used in the convolution equation, as above. In this case, no division is performed ($F = 1$), since division by zero is not defined. This generally causes the output values to be:

- zero in areas where all input values are equal (no edges)
- low in areas of low spatial frequency
- extreme in areas of high spatial frequency (high values become much higher, low values become much lower)

Therefore, a zero-sum kernel is an edge detector, which usually smooths out or zeros out areas of low spatial frequency and creates a sharp contrast where spatial frequency is high, which is at the edges between homogeneous (homogeneity is low spatial frequency) groups of pixels. The resulting image often consists of only edges and zeros. Zero-sum kernels can be biased to detect edges in a particular direction. For example, this 3 × 3 kernel is biased to the south.

-1	-1	-1
1	-2	1
1	1	1

Jensen (2005) summarized commonly spatial convolution filtering with various convolution kernels of different spatial frequencies included low-frequency filtering, high-frequency filtering and edge enhancement.

6.2.1.1 Low-frequency Filtering in the Spatial Domain

Image enhancements that de-emphasized or block the high spatial frequency detail are low-frequency or low-pass filters. The simplest *low-frequency filter* (LEF) evaluates a particular input pixel brightness value, BV_{in}, and the pixels surrounding the input pixel, and output a new brightness value, BV_{out} that is the mean of this convolution.

The primary input pixel under investigation at any one time is $BV_5 = BV_{i,j}$. The convolution of mask with all coefficients equal to 1 and the original data will result in a low-frequency filtered image, where:

$$LFF_{5,out} = \text{int}\left(\frac{\sum_{i=1}^{n=9} c_i x BV_i}{n}\right) == \text{int}\left(\frac{BV_1 + BV_2 + BV_3 + ...BV_9}{9}\right) \quad (6\text{-}8)$$

Such image smoothing is useful for removing periodic "salt and pepper" noise recorded by electronic sensing system. Some convolution masks applied as low-pass filtering (Table 6-1).

Table 6-1: Various kernels for Low-Frequency Filtering

LEF Mask	Kernel			Purposes
Mask A	1.00	1.00	1.00	For image smoothing, Salt and Peppers Remove
	1.00	1.00	1.00	
	1.00	1.00	1.00	
Mask B	0.25	0.50	0.25	To reduce blurring, unequal-weighted smoothing mask
	0.50	1	1	
	0.25	0.50	0.25	
Mask C	1.00	1.00	1.00	To reduce blurring, unequal-weighted smoothing mask
	1.00	2.00	1.00	
	1.00	1.00	1.00	

Adapted form Jensen (2005)

6.2.1.2 High-frequency Filtering in the Spatial Domain

High-pass filtering is applied to imagery to remove the slowly varying components and enhance the high-frequency local variations. One high-frequency filter ($HFF_{5,out}$) is computed by subtracting the output of the low-frequency filter ($LFF_{5,out}$) from twice the value of the original central pixel value, BV_5:

$$HFF_{5,out} = (2\ x\ BV_5) - LFF_{5,out} \quad (6\text{-}9)$$

The high-frequency filtered image will have a relatively narrows intensity histogram. It should be conducted a contrast stretching before visual interpretation. Some convolution masks applied as high-pass filtering shown in Table 6-2.

Table 6-2: Various kernels for High-Frequency Filtering

HEF Mask	Kernel			Purposes
Mask D	-1.00	-1.00	-1.00	For image accentuate or sharpen edges
	-1.00	9.00	-1.00	
	-1.00	-1.00	-1.00	
Mask E	1.00	-2.00	1.00	For image accentuate or sharpen edges
	-2.00	5.00	-2.00	
	1.00	-2.00	1.00	

Adapted form Jensen (2005)

6.2.1.3 Edge Enhancement in the Spatial Domain

For many remote sensing Earth science applications, the most valuable information that may be derived from an image is contained in the *edges* surrounding various objects of interest. *Edge enhancement* delineates these edges and makes the shapes and details comprising the image more conspicuous and perhaps easier to analyze (Jensen, 2005). Edges may be enhanced using either linear or nonlinear edge enhancement.

A. Linear Edge Enhancement.

A method of extracting edge is the application of a *directional first-difference* algorithm that approximates the first derivative between two adjacent pixels. The algorithm produces the first difference of the image input in the horizontal, vertical, and diagonal direction as:

$$Vertical = BV_{i,j} - BV_{i,j+1} + K \quad (6\text{-}10)$$

$$Horizontal = BV_{i,j} - BV_{i-1,j} + K \quad (6\text{-}11)$$

$$NE\ Diagonal = BV_{i,j} - BV_{i+1,j+1} + K \quad (6\text{-}12)$$

$$SE\ Diagonal = BV_{i,j} - BV_{i-1,j+1} + K \quad (6\text{-}13)$$

The result of the subtraction can be either negative or possible, therefore a constant, K (usually 127) is added to make all values positive and centered between 0 and 255. This causes adjacent pixels with very little difference in brightness value to obtain a brightness value of around 127 and any dramatic change between adjacent pixels to migrate away from 127 in either direction. The common kernels applying for edge enhancement were summarized in Table 6-3.

Table 6-3: Various kernels for Linear Edge Enhancement

Mask	Kernel			Purposes
Mask F	0	0	0	Emboss East For edge enhancement
	1	0	-1	
	0	0	0	
Mask G	0	0	1	Emboss NW For edge enhancement
	0	0	0	
	-1	0	0	
Mask H (North)	1	1	1	Compass Gradient mask for two dimensional, discrete differentiation directional edge enhancement
	1	-2	1	
	-1	-1	-1	
Mask I (NE)	1	1	1	Compass Gradient mask for two dimensional, discrete differentiation directional edge enhancement
	-1	-2	1	
	-1	-1	1	
Mask J (East)	-1	1	1	Compass Gradient mask for two dimensional, discrete differentiation directional edge enhancement
	-1	2	1	
	-1	1	1	
Mask K (SE)	-1	-1	1	Compass Gradient mask for two dimensional, discrete differentiation directional edge enhancement
	-1	-2	1	
	1	1	1	
Mask L (South)	-1	-1	-1	Compass Gradient mask for two dimensional, discrete differentiation directional edge enhancement
	1	-2	1	
	1	1	1	
Mask M (SW)	1	-1	-1	Compass Gradient mask for two dimensional, discrete differentiation directional edge enhancement
	1	-2	-1	
	1	1	1	
Mask N (West)	1	1	-1	Compass Gradient mask for two dimensional, discrete differentiation directional edge enhancement
	1	-2	-1	
	1	1	-1	

Table 6-3: (continued)

Mask	Kernel			Purposes
Mask O (NW)	1	1	1	Compass Gradient mask for two dimensional, discrete differentiation directional edge enhancement
	1	-2	-1	
	1	-1	-1	
Mask P (Vertical edge)	-1	0	1	Edge detection
	-1	0	1	
	-1	0	1	
Mask Q (Horizontal edge)	-1	-1	-1	Edge detection
	0	0	0	
	1	1	1	
Mask R (Diagonal)	0	1	1	Edge detection
	-1	0	1	
	-1	-1	0	
Mask S (Diagonal)	1	1	0	Edge detection
	1	0	-1	
	0	-1	-1	
Mask T (Laplacian mask)	0	-1	0	Edge enhancement
	-1	4	1	
	0	-1	0	
Mask U (Laplacian mask)	-1	-1	-1	Edge enhancement
	-1	8	-1	
	-1	-1	-1	
Mask V (Laplacian mask)	1	-2	1	Edge enhancement
	-2	4	-2	
	1	-2	1	
Mask W	1	1	1	To subtract the Laplacian edge from original image
	1	-7	1	
	1	1	1	
Mask X	0	-1	0	To create original image back from the Laplacian edge enhancement image
	-1	5	-1	
	0	-1	0	

Table 6-3: (continued)

Mask	Kernel					Purposes
Mask X		0	-1	0		To create original image back from the Laplacian edge enhancement image
		-1	5	-1		
		0	-1	0		
Mask Y	0	0	-1	0	0	To create original image back from the Laplacian edge enhancement image
	0	-1	-2	-1	0	
	-1	-2	17	-2	-1	
	0	-1	-2	-1	0	
	0	0	-1	0	0	

Adapted form Jensen (2005)

B. Nonlinear Edge Enhancement

Nonlinear edge enhancements are performed using nonlinear combinations of pixels. Many algorithms are applied using 2 x2 or 3 x 3 kernels (Jensen, 2005). The *Sobel edge detector* is based on the notation of 3 X 3 windows and is computed according to the relationship:

$$Sobel_{5,out} = \sqrt{X^2 + Y^2} \qquad (6\text{-}14)$$

where:

$X = (BV_3 + 2BV_6 + BV_9) - (BV_1 + 2BV_4 + BV_7)$ and

$Y = (BV_1 + 2BV_2 + BV_3) - (BV_7 + 2BV_8 + BV_9)$.

The Sobel operator may also be computed by simultaneously applying the following 3x3 template across the image:

$$X = \begin{matrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{matrix} \qquad Y = \begin{matrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{matrix}$$

This procedure detects horizontal, vertical, and diagonal edges.

Also, Robert's edge detector can be computed by simultaneously applying the following 3x3 template across the image:

$$X = \begin{matrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{matrix} \qquad Y = \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{matrix}$$

Furthermore, detection of line can applied as well as edge with various masks in various directions as shown in Table 6-4.

Example of various spatial convolution filtering was shown in Figure 6-15 and 6-16.

Table 6-4: Various kernels for Line Detection

Mask	Kernel			Purposes
Mask AA	-1	-1	-1	Line detection: E – W
	2	2	2	
	-1	-1	-1	
Mask BB	-1	-1	2	Line detection: NE – SW
	-1	2	-1	
	2	-1	-1	
Mask CC	-1	2	-1	Line detection: N-S
	-1	2	-1	
	-1	2	-1	
Mask DD	2	-1	-1	Line detection: NW-SE
	-1	2	-1	
	-1	-1	2	

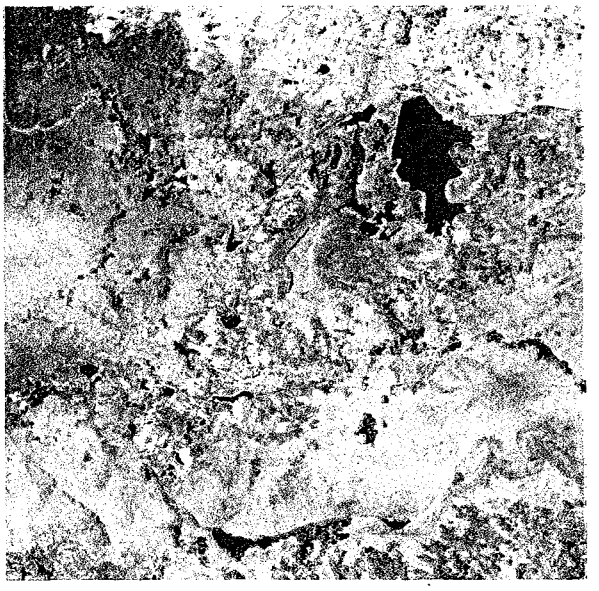
Adapted form Jensen (2005)

In addition, spatial convolution filtering that was categorized into three types included linear, statistical and gradient filters was summarized by Schowengerdt (1997) as shown in Table 6-5.

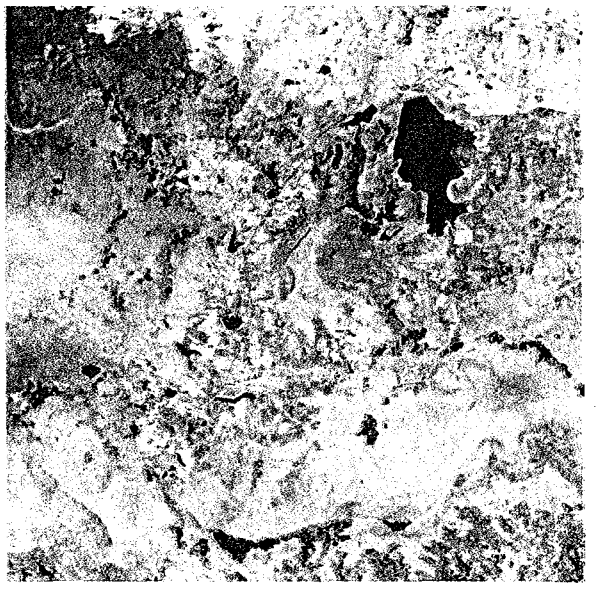
Table 6-5: Catalog of local convolution filtering

Type	Output	Examples	Application
Linear	Weight sum	• Low-pass filter (LPF) • High-pass filter (LPF) • High-boost filters (HBF)	• Enhancement • Sensor simulation
Statistical	Given statistic	• Minimum • Maximum • Median • Mode • Standard deviation	• Noise removal • Feature extraction • Signal-to-Noise Ratio (SNR) measurement
Gradient	Vector gradient	• Sobel • Roberts	• Edge detection

From: Schowengerdt (1997)

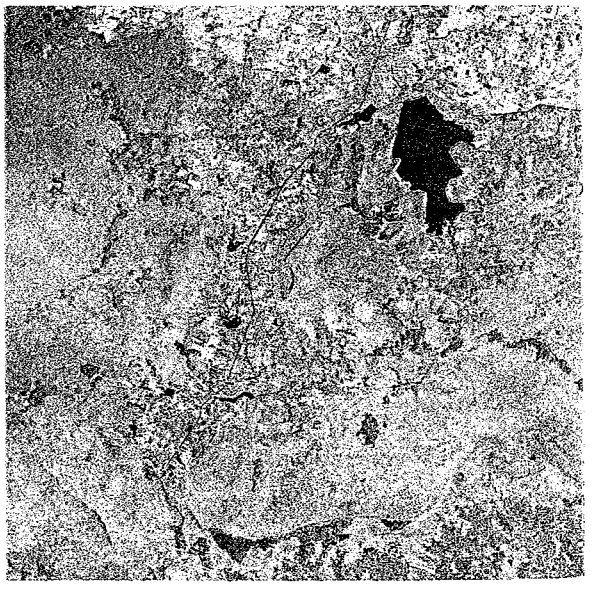


(a) Original Contrast Stretching



(b) Median Filtering

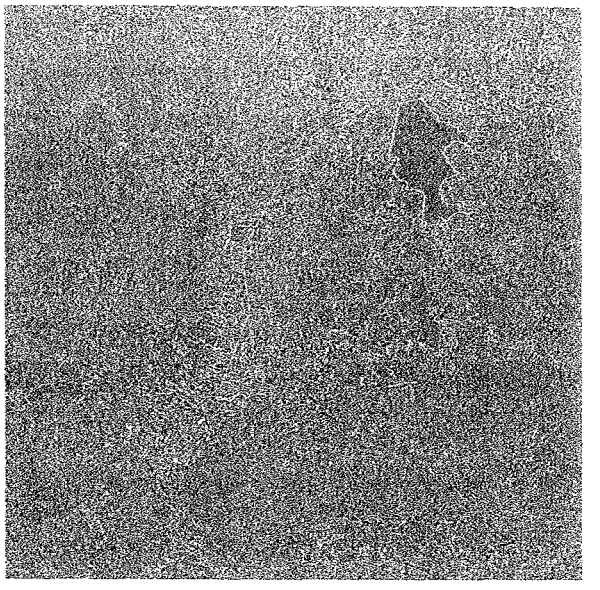
(c) Low-frequency filtering



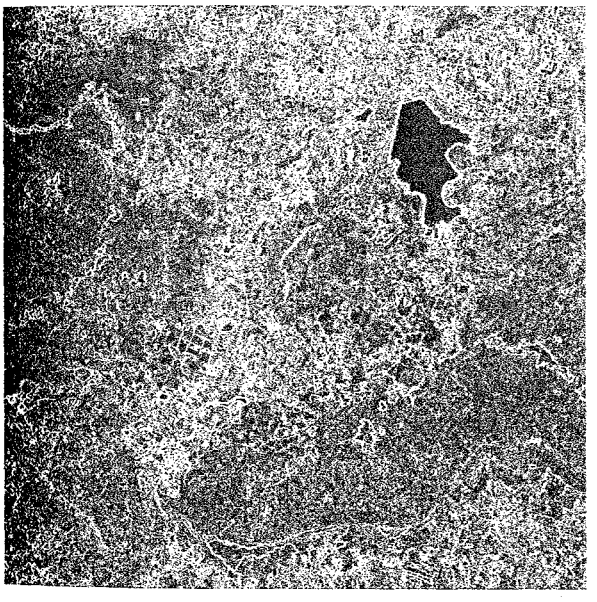
(d) High-frequency filtering

Figure 6-15: Application of various convolution masks and logic to Landsat TM Band 4 data: Date 6 March 2005, Non Din Deang, Buriram, Thailand: (a) Original contrast stretched data (b) Median filtering (c) Low-frequency filtering and (d) High-frequency filtering.

(a) Original Contrast Stretching



(b) Laplacian edge Detection



(c) Sobel's edge detection

(d) Prewitt's edge detection

Figure 6-16: Application of various convolution masks and logic to Landsat TM Band 4 data, Date 6 March 2005, Non Din Deang, Buriram, Thailand: (a) Original contrast stretched data (b) Laplacian edge detection (c) Sobel's edge detection and (d) Prewitt's edge detection.

6.2.2 Fourier Transformation

Fourier analysis is a mathematical technique for separating and image into its various spatial frequency components. Fourier theorem states that any function *f(x)* can be represented by a summation of a series of sinusoidal terms of varying spatial frequencies (Jensen, 2005). For example, a line of pixels with a high spatial frequency gray scale pattern might be represented in terms of a single coefficient multiplied by a sin(x) function. High spatial frequencies are those that represent frequent gray scale changes in a short pixel distance. Low spatial frequencies represent infrequent gray scale changes that occur gradually over a relatively large number of pixel distances. A more complicated function, *f(x)*, might have to be represented by many sine and cosine terms with their associated coefficients.

Figure 6-17 shows how a function *f(x)* can be represented as a linear combination of sine and cosine. In this example the function is a square wave whose cosine coefficients are zero leaving only sine terms. The first three terms of the Fourier series are plotted in the upper right graph and the plot of the sum is shown below it. After nine iterations, the Fourier series is approaching the original function.

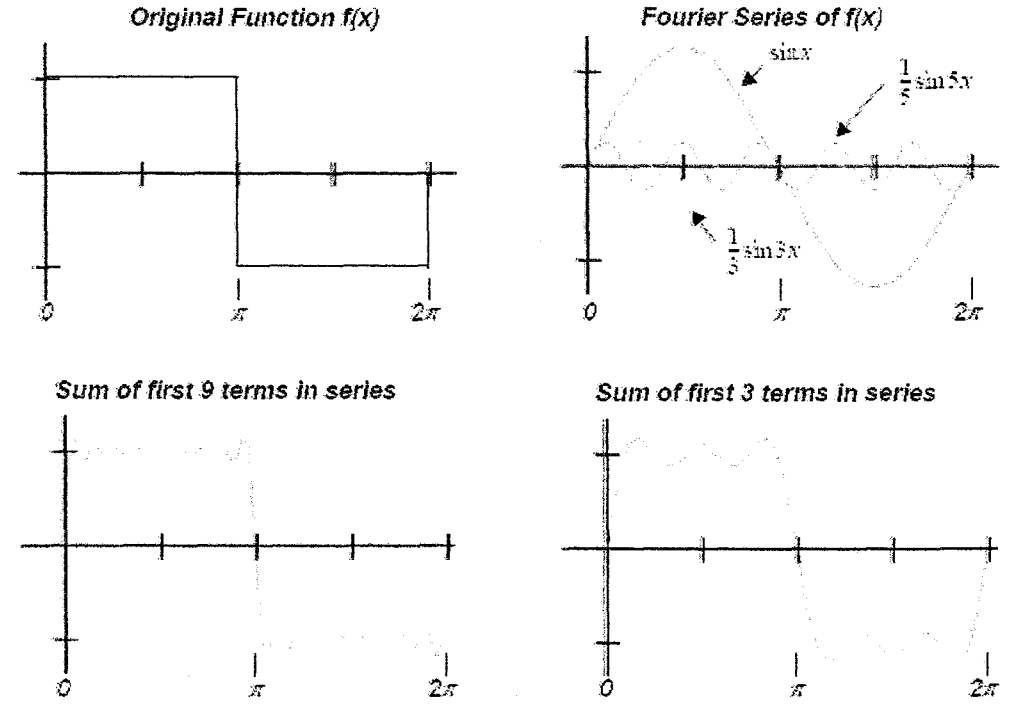


Figure 6-17: One-Dimensional Fourier Analysis (From ERDAS, 2002).

To use Fourier analysis in digital image processing, two-dimensional discrete Fourier transform will be firstly calculate the Fast Fourier Transform (FFT) of image as following:

$$F(u,v)=\frac{1}{NM}\sum_{x=0}^{N-1}\sum_{y=0}^{M-1}f(x,y)e^{-2\pi i(\frac{ux}{N}+\frac{vy}{M})} \quad (6\text{-}15)$$

Where:

M = the number of pixels horizontally

N = the number of pixels vertically

u,v = spatial frequency variables

e = 2.71828, the natural logarithm base

i = the imaginary component of a complex number

Secondly, the image can be reconstructed using the Inverse Fast Fourier Transform (IFFT) by following equation:

$$f(x,y)=\sum_{u=0}^{N-1}\sum_{v=0}^{M-1}F(u,v)e^{-2\pi i(\frac{ux}{N}+\frac{vy}{M})} \quad (6\text{-}16)$$

It contains the spatial frequency information of the original image *f(x, y)* and is called the frequency spectrum with containing of complex function ($i=\sqrt{-1}$). It can write this complex function as the sum of a real part and an imagery part as:

$$F(u,v)=R(u,v)+iI(u,v) \quad (6\text{-}17)$$

which is equivalent to

$$F(u,v)=|F(u,v)|e^{i\phi(u,v)} \quad (6\text{-}18)$$

Where $|F(u,v)|$ is a real function, and $|F(u,v)|=\sqrt{R(u,v)^2+I(u,v)^2}$.

$|F(u,v)|$ is called the magnitude of Fourier Transform. It represents the magnitude and the direction of the different frequency components in image *f(x, y)*. The variables ϕ in Equation 6-x represents phase information in the image *f(x, y)*. It will use in recovery of image (Jensen, 2005).

Figure 6-18 shows example of three varieties of remote sensed data and their Fourier transform and example of stationary periodic noise and their Fourier transform in Figure 6-19.

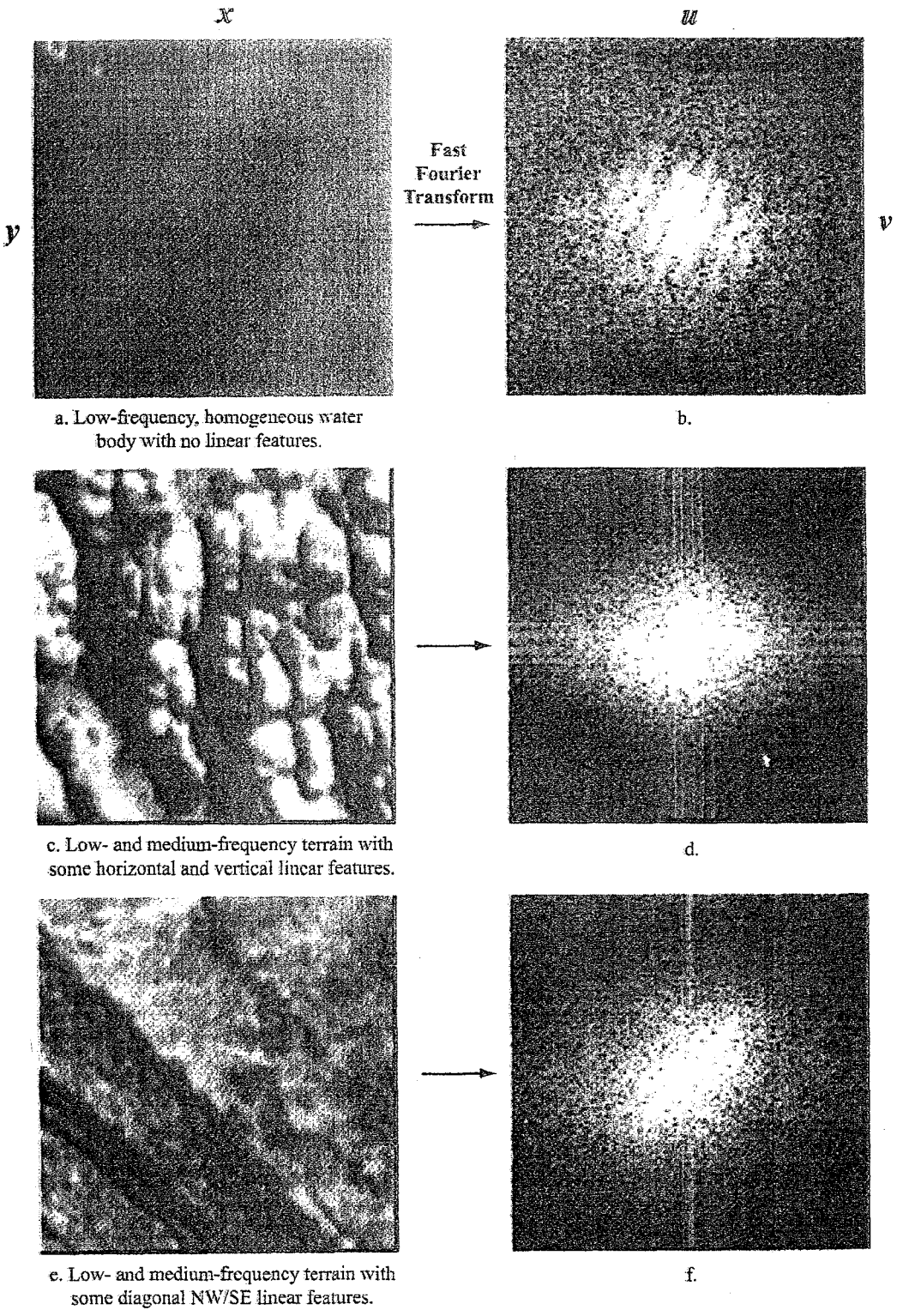
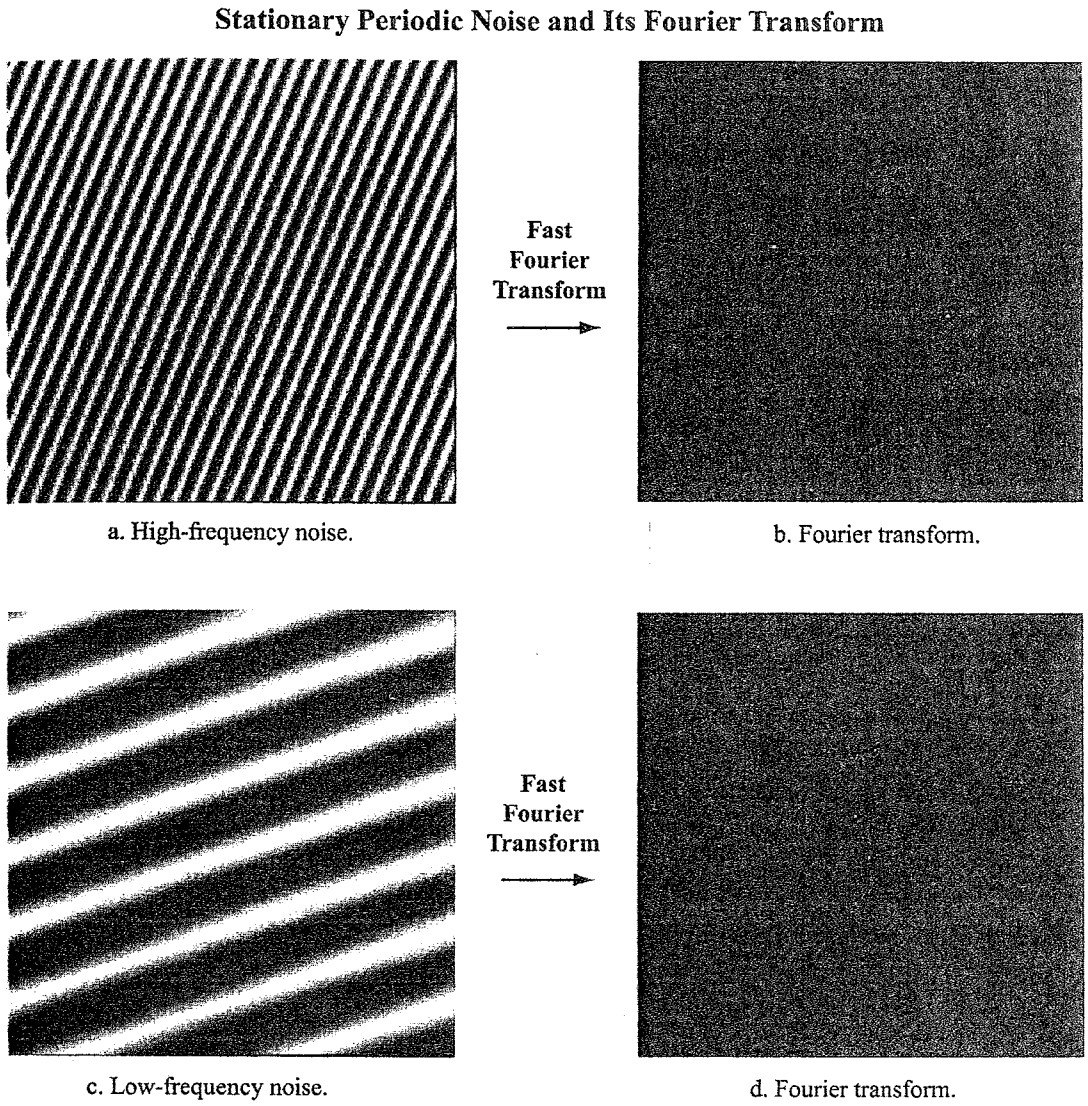


Figure 6-18: Application of Fourier transform to the three different sub-images (From Jensen, 2005)



ɐ 8-29 Two examples of stationary periodic noise and their Fourier transforms.

Figure 6-19: Two examples of stationary periodic noise and their Fourier transforms (From Jensen, 2005).

6.2.2.1 Spatial Filter in Frequency Domain

Using the Fourier transform, we can manipulate directly the frequency information of the image. The manipulate can be performed by multiplying the Fourier transform of the original image by a mask image called a **frequency domain filter**, which will block or weaken certain frequency components by making the values of certain parts of the frequency spectrum become smaller or even zero. Then, we can compute the inverse Fourier transform of the manipulated frequency spectrum to obtain a filtered image in the spatial domain. Numerous algorithms are available for computing the Fast Fourier transform (FFT) and the inverse Fast Fourier transform (IFFT).

Practically, spatial filtering in the frequency domain generally involves computing the FFT of the original image, multiplying the FFT of a convolution mask of the analyst's choice (e.g., low pass or high pass) with the FFT, and inverting the resultant image with IFFT (Jensen, 2005) that is,

$f(x,y)$ FFT $\rightarrow F(u,v) \rightarrow F(u,v)\ G(u,v) \rightarrow F'(u,v)$ IFFT $\rightarrow f'(x,y)$

The convolution theorem states that the convolution of two images is equivalent to the multiplication of their Fourier transform. If

$$f'(x,y) = f(x,y) \overset{*}{} g(x,y) \tag{6-19}$$

where:

*	represents the operation of convolution,
f(x,y)	is the original image and;
g(x,y)	is a convolution mask filter.

Then,

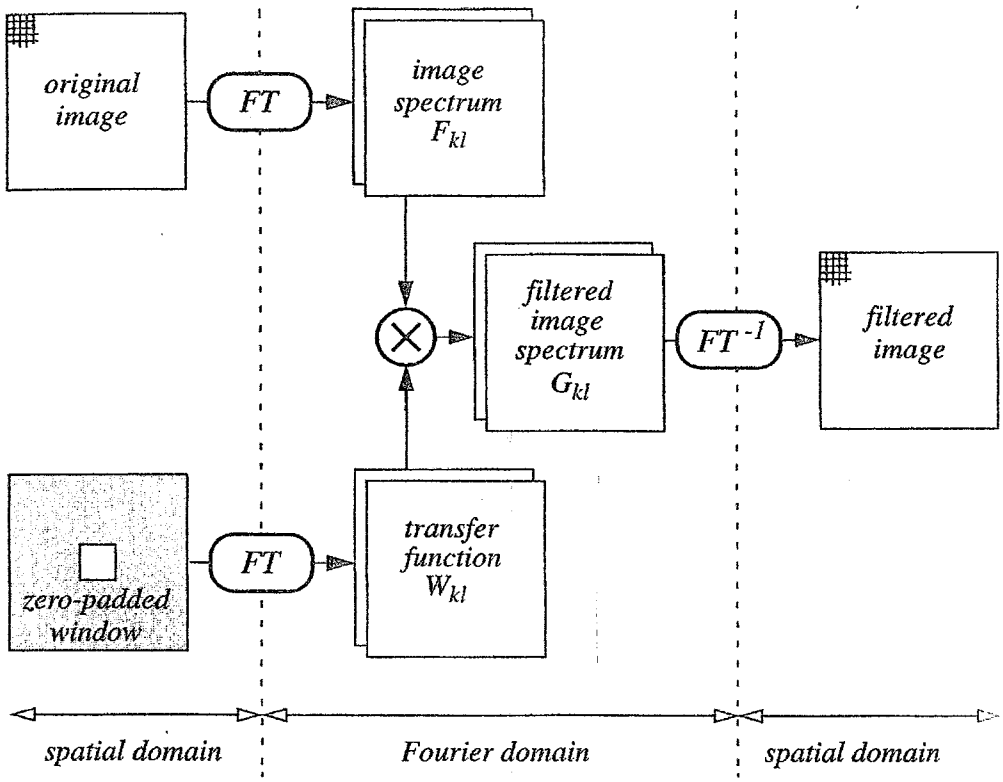
$$F'(u,v) = F(u,v)G(u,v) \tag{6-20}$$

where F', F and G are Fourier transform of f', f, and g, respectively.

The multiplication of Fourier transform F (u,v) and G (u,v) results in a new Fourier transform, F'(u,v). Computing the Inverse Fourier transformation yields f'(x,y), a filtered version of the original image. Thus, spatial filtering can be performed in both the spatial and frequency domain.

To use Fourier transforms for a spatial domain convolution, we must take the Fourier transform of the original image and the window weighting function, multiply the image spectra and the transfer function, which yields the filtered image spectrum, and then take the inverse of that product to obtain the filtered image in the spatial domain (Schowengerdt, 1997). This sequence is depicted in Figure 6-20.

Two examples of spatial filtering in frequency domain using a Fourier transform were shown in Figure 6-21 and Figure 6-22.



Note: The Fourier domain arrays are shown as double because the data are complex in that domain

Figure 6-20: A filtering algorithm that uses the Fourier transform to compute a spatial domain convolution (Schowengerdt, 1997).

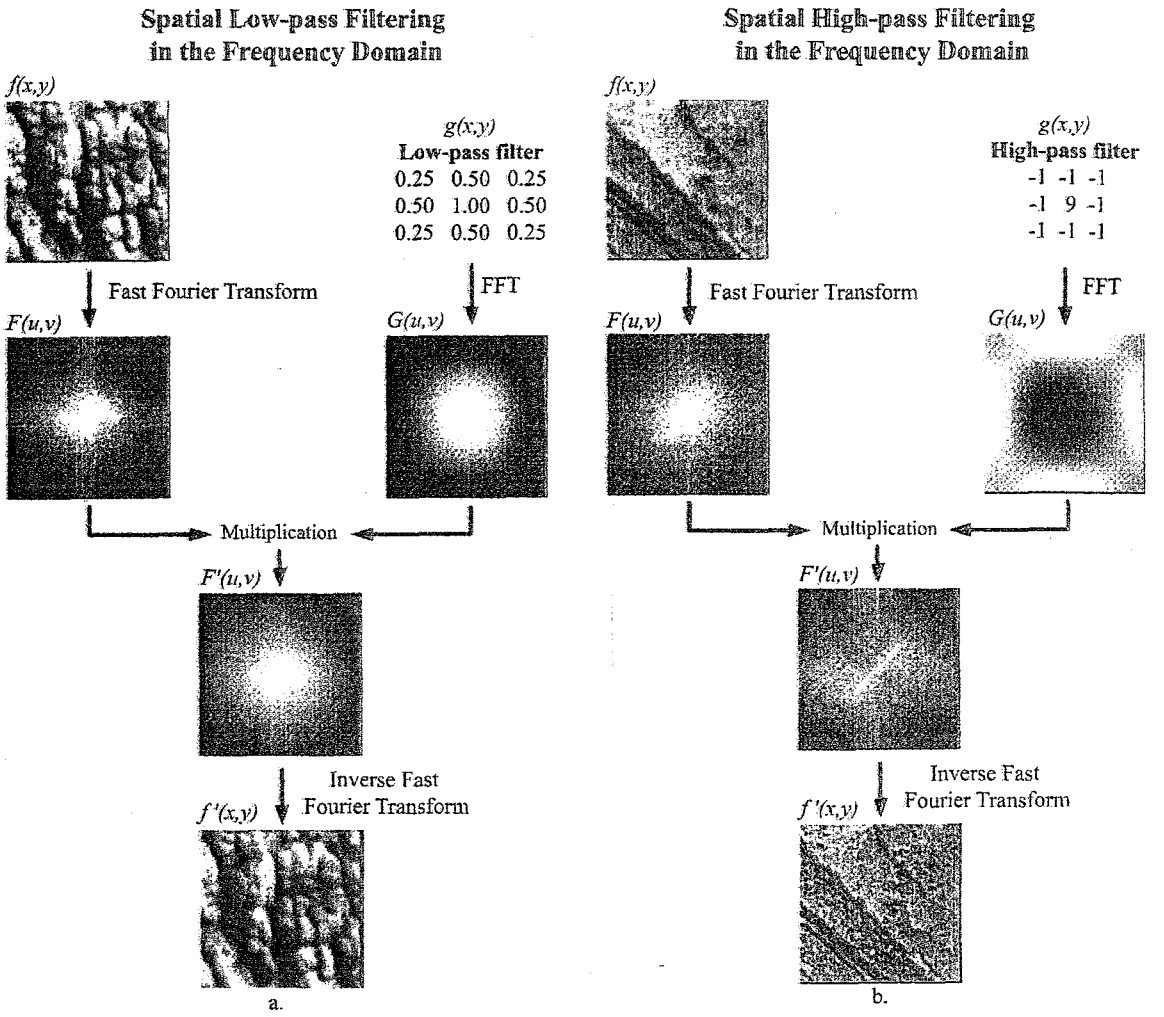
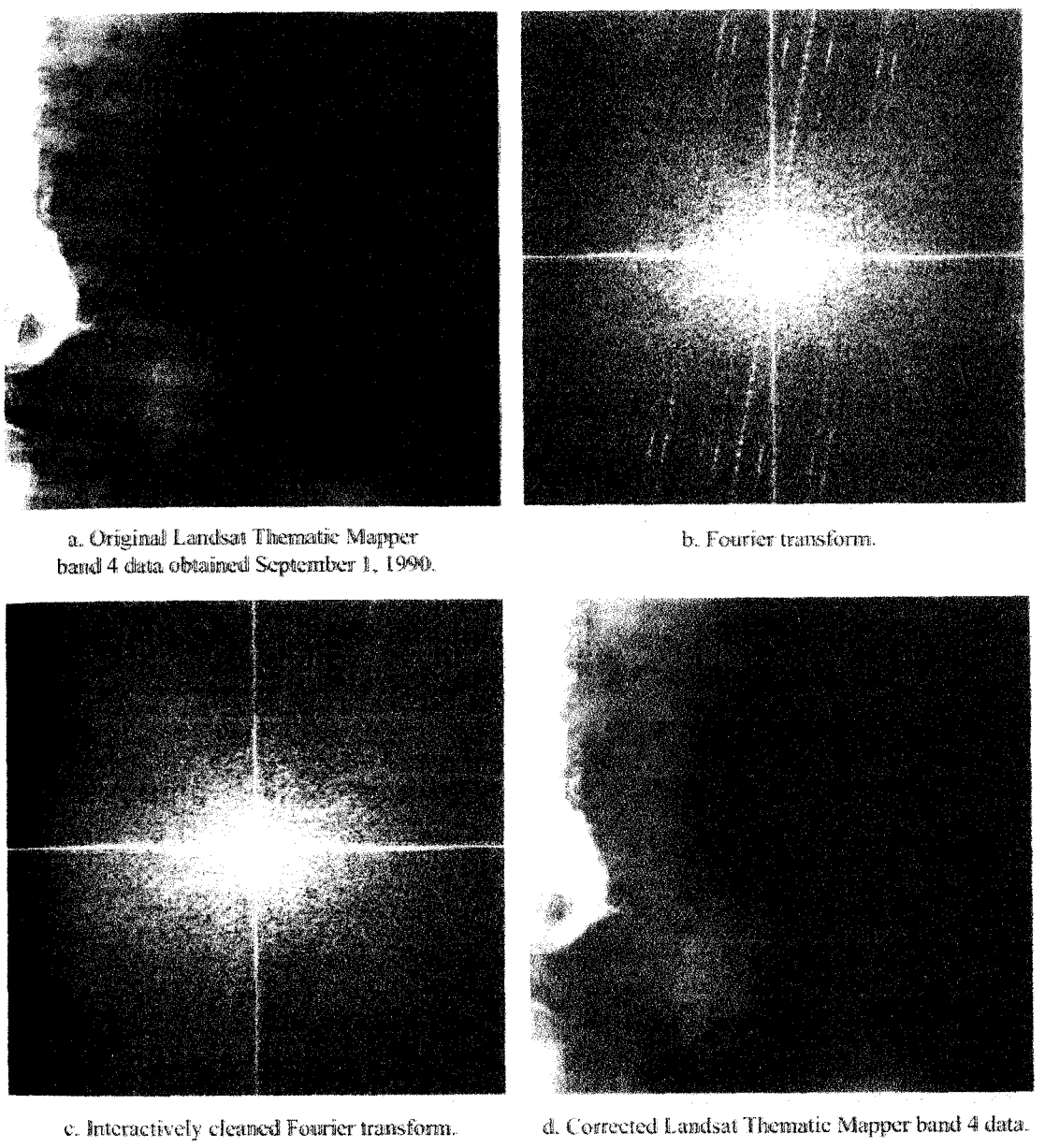


Figure 6-21: Spatial Filter in Frequency Domain using a Fourier transform (a) Spatial low-pass and (b) Spatial high-pass filtering. (From, Jensen, 2005)



a. Original Landsat Thematic Mapper band 4 data obtained September 1, 1990.

b. Fourier transform.

c. Interactively cleaned Fourier transform.

d. Corrected Landsat Thematic Mapper band 4 data.

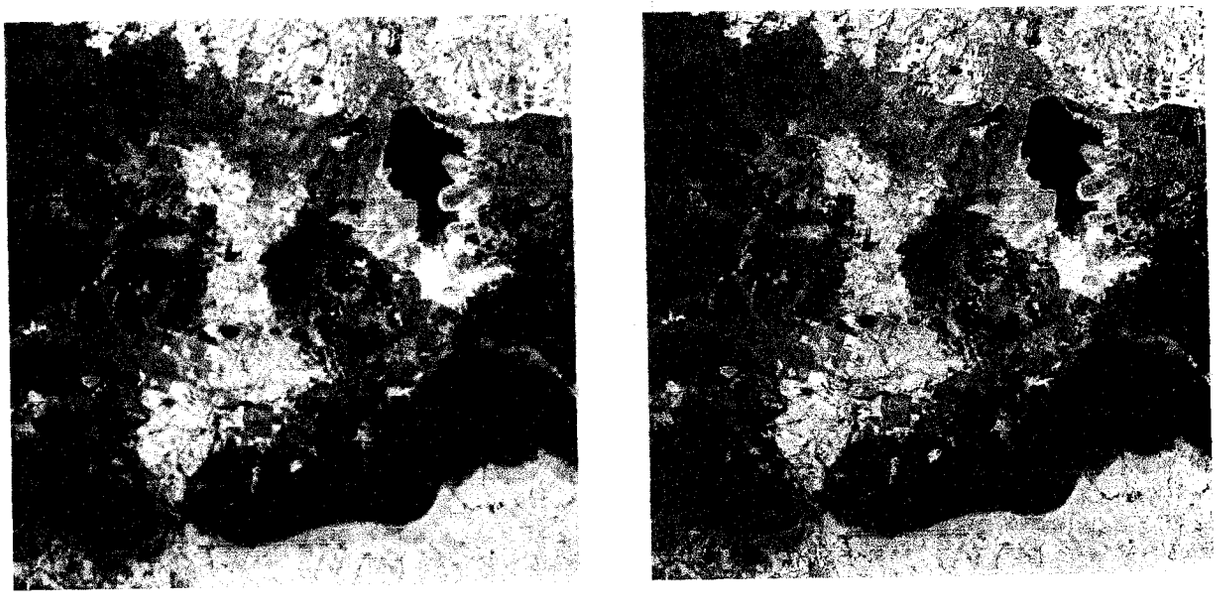
Figure 6-22: Application of Fourier transform to a portion of Landsat TM band 4 data of Al Jabail, Saudi Arabia (From Jensen, 2005).

6.2.3 CRISP

The Crisp filter sharpens the overall scene luminance without distorting the interband variance content of the image. This is a useful enhancement if the image is blurred due to atmospheric haze, rapid sensor motion, or a broad point spread function of the sensor (ERDAS, 2002). The algorithm used for this function is:

- Calculate principal components of multiband input image.
- Convolve PC-1 with summary filter.
- Retransform to RGB space.

The logic of the algorithm is that the first principal component (PC-1) of an image is assumed to contain the overall scene luminance. The other PCs represent intra-scene variance. Thus, you can sharpen only PC-1 and then reverse the principal components calculation to reconstruct the original image. Luminance is sharpened, but variance is retained. Comparison of original Landsat TM band 4,5, 3 (RGB) and Crisp Filtering was shown in Figure 6-23.



Original False Color Composite Landsat-TM Band 4,5,3 (RGB)

CRISP Filtering Imagery

Figure 6-23: Comparison of original Landsat TM composite data (Band 4, 5, and 3: RGB): Date 6 March 2005, Non Din Deang, Buriram, Thailand and Crisp Filtering

6.2.4 Resolution Merge

The resolution of a specific sensor can refer to radiometric, spatial, spectral, or temporal resolution. Landsat TM sensors have seven bands with a spatial resolution of 28.5 m. SPOT panchromatic has one broad band with very good spatial resolution—10 m. Combining these two images to yield a seven-band data set with 10 m resolution provides the best characteristics of both sensors.

ERDAS (2002) reviewed number of models have been suggested to achieve this image merge included:

- Welch and Ehlers (1987) used forward-reverse RGB to IHS transforms, replacing I (from transformed TM data) with the SPOT panchromatic image. However, this technique is limited to three bands (R, G, B).
- Chavez (Chavez et al, 1991), among others, uses the forward-reverse principal components transforms with the SPOT image, replacing PC-1.

In the above two techniques, it is assumed that the intensity component (PC-1 or I) is spectrally equivalent to the SPOT panchromatic image, and that all the spectral information is contained in the other PCs or in H and S. Since SPOT data do not cover the full spectral range that TM data do, this assumption does not strictly hold. It is unacceptable to resample the thermal band (TM6) based on the visible (SPOT panchromatic) image.

Another technique (Schowengerdt, 1980) combines a high frequency image derived from the high spatial resolution data (i.e., SPOT panchromatic) additively with the high spectral resolution Landsat TM image.

Three common techniques for resolution merge includes (a) Principal component merge (b) Multiplication and (c) Brovey Transform based on ERDAS (2002) here are reviewed in detail.

6.2.4.1 Principal Components Merge

Because a major goal of this merge is to retain the spectral information of the six TM bands 1 - 5, 7), this algorithm is mathematically rigorous. It is assumed that:

• PC-1 contains only overall scene luminance; all interband variation is contained in the other 5 PCs, and

• scene luminance in the SWIR bands is identical to visible scene luminance.

With the above assumptions, the forward transform into PCs is made. PC-1 is removed and its numerical range (min to max) is determined. The high spatial resolution image is then remapped so that its histogram shape is kept constant, but it is in the same numerical range as PC-1. It is then substituted for PC-1 and the reverse transform is applied. This remapping is done so

that the mathematics of the reverse transform do not distort the thematic information (Welch and Ehlers, 1987).

6.2.4.2 Multiplicative Algorithm

The second technique in the Image Interpreter uses a simple multiplicative algorithm:

(DNTM1) (DNSPOT) = DNnew TM1

The algorithm is derived from the four component technique of (Crippen, 1989). In this paper, it is argued that of the four possible arithmetic methods to incorporate an intensity image into a chromatic image (addition, subtraction, division, and multiplication), only multiplication is unlikely to distort the color.

However, in his study Crippen first removed the intensity component via band ratios, spectral indices, or PC transform. The algorithm shown above operates on the original image. The result is an increased presence of the intensity component. For many applications, this is desirable. People involved in urban or suburban studies, city planning, and utilities routing often want roads and cultural features (which tend toward high reflection) to be pronounced in the image.

6.2.4.3 Brovey Transform

In the Brovey Transform method, three bands are used according to the following formula:

[DNB1 / DNB1 + DNB2 + DNB3] × [DNhigh res. image] = DNB1_new

[DNB2 / DNB1 + DNB2 + DNB3] × [DNhigh res. image] = DNB2_new

[DNB3 / DNB1 + DNB2 + DNB3] × [DNhigh res. image] = DNB3_new

Where:

B(n) = band (number)

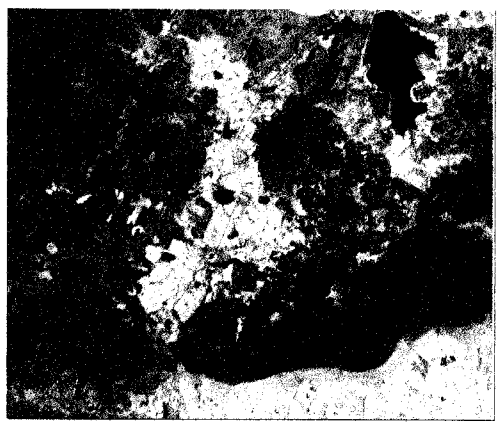
The Brovey Transform was developed to visually increase contrast in the low and high ends of an image's histogram (i.e., to provide contrast in shadows, water and high reflectance areas such as urban features). Consequently, the Brovey Transform should not be used if preserving the original scene radiometry is important. However, it is good for producing RGB images with a higher degree of contrast in the low and high ends of the image histogram and for producing visually appealing images.

Since the Brovey Transform is intended to produce RGB images, only three bands at a time should be merged from the input multispectral scene, such as bands 3, 2, 1 from a SPOT or Landsat TM image or 4, 3, 2 from a Landsat TM image. The resulting merged image should then be displayed with bands 1, 2, 3 to RGB.

Example of resolution merge between SPOT with 10 m spatial resolution and Landsat-TM data with 30 m spatial resolution was shown in Figure 6-24.

(a) Principal Components Merge

(b) Multiplicative Algorithm

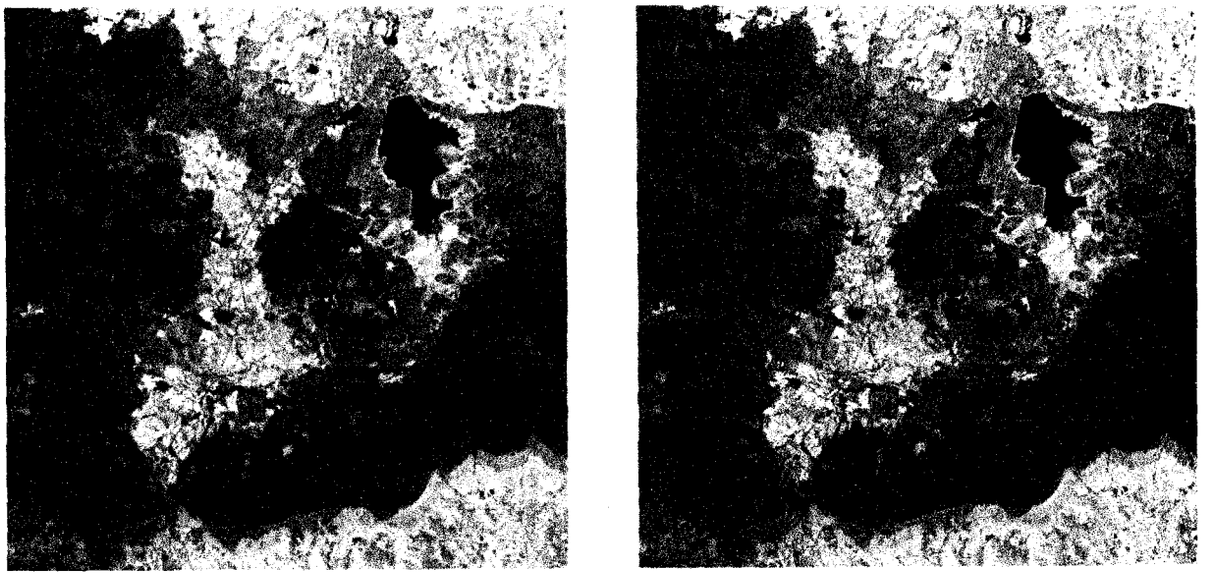


(c) Brovey Transform

Figure 6-24: Resolution Merge between SPOT data (25 Jan 2006) and Landsat TM (9 Mar 2006) of Non Din Deang, Buriram, Thailand: (a) Principal Components Merge, (b) Multiplicative algorithm and (c) Brovey Transform

6.2.5 Adaptive Filter

Adaptive Filter Contrast enhancement (image stretching) is a widely applicable standard image processing technique (ERDAS, 2002). However, even adjustable stretches like the piecewise linear stretch act on the scene globally. There are many circumstances where this is not the optimum approach. For example, coastal studies where much of the water detail is spread through a very low DN range and the land detail is spread through a much higher DN range would be such a circumstance. In these cases, a filter that adapts the stretch to the region of interest (the area within the moving window) would produce a better enhancement. Adaptive filters attempt to achieve this (Fahnestock and Schowengerdt, 1983; Peli and Lim, 1982; Schwartz and Soha, 1977).



Original False Color Composite Landsat-TM Band 4,5,3 (RGB)

Adaptive Filtering Imagery

Figure 6-25: Comparison of original Landsat TM composite data (Band 4, 5, and 3: RGB): Date 6 March 2005, Non Din Deang, Buriram, Thailand and Adaptive Filtering

6.2.6 Texture Transform

Texture is the degree of roughness or smoothness of an object or surface in the image. Pattern is the regular repetition of a tonal arrangement across an image. In practice, texture is a local phenomenon and as a consequence it can be measured within the local environs of the pixel. Thus, many measures of texture operate within a window around a pixel, and the derived textural values are assigned to the central pixel of that window.

There are several standard approaches to automatic texture classification, including texture features based on first-and second order gray-level statistics and on the Fourier power spectrum and measures based on fractals. Jensen (2005) mentioned that several studies have concluded that the use of the Fourier transform for texture analysis generally yields poor results.

A. First-Order Statistics in the Spatial Domain

One class of picture properties that can be used for texture synthesis is first-order statistics of local areas such as means (average), variance, standard deviation, and entropy. Typical algorithms include:

$$AVE = \frac{1}{w} \sum_{i=0}^{quant_k} i \times f_i \tag{6-21}$$

$$STD = \sqrt{\frac{1}{w} \sum_{i=0}^{quant_k} (i - ave)^2 \times f_i} \tag{6-22}$$

$$ENT_1 = \sum_{i=0}^{quant_k} \frac{f_i}{w} ln \frac{f_i}{w} \tag{6-23}$$

where:

f_i = frequency of gray level i occurring in a pixel window

$quant_k$ = quantization level of band k

w = total number of pixels in a window

B. Second-Order Statistics in the Spatial Domain

Jensen (2005) mentioned that a suite of very useful texture measures was originally developed by Haralick and associates. The higher-order set of texture measures is based on brightness value spatial-dependency gray-level co-occurrence matrices (GLCM). The GLCM-derived

texture transformations have been widely adopted by the remote sensing community and are often used as an additional feature in multispectral classification.

Basically, if $c = (\Delta x, \Delta y)$ is considered a vector in the (x, y) image plane, for any such vector and for any image f(x, y) it is possible to compute the joint probability density of the pairs of brightness values that occur at pairs of the points separated by c. If the brightness values in the image can take upon themselves any value from 0 to the highest quantization level in the image, this joint density takes the from of an array h_c, where $h_c(i,j)$ is the probability that the pairs of brightness values (i, j) occur at separation c. This array h_c is qaunt$_k$ by quant$_k$ in size. It is easy to compute the h_c array for $f(x, y)$, where Δx and Δy are simply counting the number of times each pair of brightness values occurs at separation c (Δx and Δy) in the image (Jensen, 2005). For example, consider the following image that has just lines and five columns and contains brightness values ranging from only 0 to 3:

Original image =

0	1	1	2	3
0	0	2	3	3
0	1	2	2	3
1	2	3	2	2
2	2	3	3	2

If (Δx and Δy) = (1,0), then these numbers are represented by the brightness values spatial dependency matrix h_c :

h_c =

	0	1	2	3
0	1	2	1	0
1	0	1	3	0
2	0	0	3	5
3	0	0	2	2

Where the entry in row i and column j of this matrix is the number of times brightness value i occurs to the left of brightness value j. For example, brightness value 1 is to the left of brightness value 2 a total of three times in this simplified image [i.e., $h_c(1,2) = 3$]. It is assumed that all textural information is contained in the brightness value spatial dependency matrices that are developed for angle of 0°, 45°, 90°, and 135° (Figure 6-26). Generally, the greater the number found in the diagonal of the gray-level co-occurrence matrices, the more homogeneous the texture is for that part of the image being analyzed (Jensen, 2005).

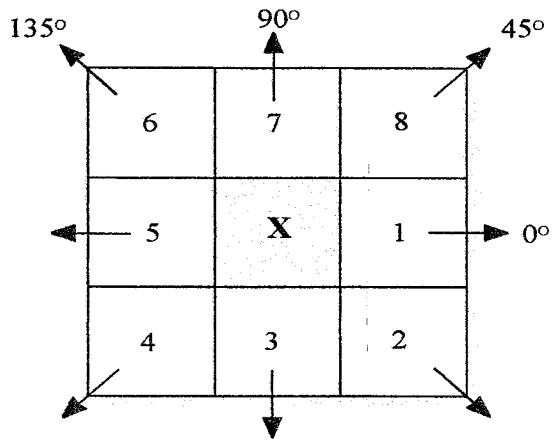


Figure 6-26: The eight nearest neighbors of pixel X according to angle ϕ used in the creator of spatial-dependency matrices for the measurement of image texture (From Jensen, 2005).

There are a variety of measures that can be used to extract useful textural information from the h_c matrices. Four of the more widely used include the angular second moment (ASM), contrast (CON), correlation (COR), entropy (ENT2) and homogeneity (HOM):

$$ASM = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} h_c(i,j)^2 \quad (6\text{-}24)$$

$$CON = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} (i-j)^2 \times h_c(i,j)^2 \quad (6\text{-}25)$$

$$COR = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} \frac{(i-\mu)(j-\mu)h_c(i,j)^2}{\sigma^2} \quad (6\text{-}26)$$

$$ENT_2 = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} h_c(i,j) \times log[h_c(i,j)] \quad (6\text{-}27)$$

$$HOM = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} \frac{1}{1+(i-j)^2} h_c(i,j)^2 \quad (6\text{-}28)$$

where:

quantk = quantization level of band *k* (e.g. 28 = 0 to 255)

$h_c(i, j)$ = the (i, j)th entry in one of the angular brightness value spatial-dependency matrices, and

$$\mu = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} i \times h_c(i,j) \quad (6\text{-}29)$$

$$\sigma^2 = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} (i-\mu)^2 \times h_c(i,j) \quad (6\text{-}30)$$

During computation, four brightness value spatial-dependency matrices (0°,45°,90° and 135°) are derived for each pixel based on neighboring pixel values. The average of these four measures is normally output as the texture value for the pixel under consideration.

To use a GLCM-derived texture measure, the analyst has to make important decisions, including (Franklin et al., 2001):

- the texture measures;
- window size;
- input channel;
- quantization level of the input data;
- the spatial component.

Furthermore, McCloy (2006) mentioned that there are three basic approaches to the measurement of texture: statistical, structural and spectral. Statistical approaches derive the moments within the window used about the pixel. Structural approaches seek to detect the effect of small-scale patterns in the image space. Spectral techniques utilize the properties of the Fourier transform to detect regular patterns within the image space.

A. Statistical Techniques

The simplest techniques to measure texture are to derive the second, third and fourth moments within the window about the pixel where:

$$\text{Second moment = variance} = \sum(g(i,j)-m)^2 \times p(g(i,j)) = \sigma^2 \qquad (6\text{-}31)$$

where g(i,j) = picture function, m = mean value and p() = probability of occurrence.

The variance is an important measure of the variability in the data in the window or sample area, with σ^2 - > 0.0 for smooth areas and with increasing values for areas of increasing roughness. Sometimes this measure can be normalized by:

$$R = (\sigma - \sigma_{min})/(\sigma_{max} + \sigma_{min}) \qquad (6\text{-}32)$$

where R takes a value near zero for smooth areas and approaches one for rough areas, and σmin and σmax are the minimum and maximum local variance values found in the image.

The third moment is a measure of the skewness of the histogram:

$$Skewness = \sum(g(i,j)-m)^3 \times p(g(i,j)) \qquad (6\text{-}33)$$

If the data in the window is clustered about a value, and there is a tail of a few high, or low values then the histogram will be skewed.

Then fourth moment is a measure of the kurtosis or peakedness of the histogram:

$$Kurtosis = \sum(g(i,j)-m)^4 \times p(g(i,j)) \qquad (6\text{-}34)$$

Measures of texture that are based on these statistics suffer the disadvantage that they provide no information on patterns in the data.

B. Structural Methods

Structural measures of texture attempt to provide some information on these characteristics. The most common structural method is the construction of co-occurrence matrices for each pixel using the data in a window around that pixel. Consider the image data shown in Figure 6-27. Choose a position operator that will detect the features of interest.

2	3	3	1	0
0	2	3	1	0
0	0	3	3	2
2	0	1	3	3
3	2	0	1	3

Figure 6-27: A 5 X 5 window of pixels and their image values (From McCloy, 2006).

Typical position operators are given by McCloy (2006) as following.

1. Detection of Edge at +45°. For this position operator, create an array of rows I and columns j such that a location [i, j] in the array is the sum of the number of occurrences of a pixel of value i being located one down and one to the right of a pixel of value j. Thus the first location in this array, [0, 0], is the count of the number of occurrence of 0 in the window down one line and to the right by one pixel from another value of 0. In the data shown in Figure 6-27 this location will have a value of 3. The full array for this data and with this position operator is:

3	1	0	0
1	1	0	1
0	1	2	0
0	0	1	5

2. Detection of edges at -45°. This position operator is similar to 1 above, but the first pixel position is one down and to the right of the second pixel. The values in the array for this position operator and this window of data are:

0	1	1	2
0	0	1	2
1	1	0	1
2	2	1	0

3. Detection of horizontal edges. For this position operator, create an array of rows i and columns j such that a location [i, j] in the array is the sum of the number of occurrences of a pixel of value i being one down from a pixel of value j. The values in this array must sum to 20:

3	0	3	0
1	1	0	1
2	0	0	2
0	2	1	4

4. Detection of vertical edges. This position operator is similar to 3 above, but value i is to be to the right of value j. The result of this analysis for this window of data is:

1	2	2	0
2	0	0	2
2	0	0	2
0	2	2	3

These arrays are then divided by this total, 16 for 1 and 2, and 30 for 3 and 4, to give the co-occurrence arrays C[i, j].

For the first position operator:

0.1875	0.0625	0	0
0.0625	0.0625	0	0.0625
0	0.0625	0.125	0
0	0	0.0625	0.3125

For the second position operator:

0	0.0625	0.0625	0.125
0	0	0.0625	0.125
0.0625	0.0625	0	0.0625
0.125	0.125	0.0625	0

For the third position operator:

0.15	0	0.15	0
0.05	0.05	0	0.05
0.1	0	0	0.1
0	0.1	0.05	0.2

And the fourth position operator:

0.05	0.1	0.1	0
0.1	0	0	0.1
0.1	0	0	0.1
0	0.1	0.1	0.15

The co-occurrence matrix can then be used to derive various indices, the usual ones being:

1. Maximum probability $max[c(i,j)]$
2. Element-difference moment of order k $\sum\sum(i-j)\times C(i,j)$
3. Inverse element-difference moment of order k $\sum\sum C(i,j)/(i-j)^k$
4. Entropy $-\sum\sum C(i,j)\times log(C(i,j))$
5. Uniformity $\sum\sum C(i,j)\times C(i,j)$

C. Spectral Techniques

These approaches to texture analysis depend on the Fourier transform or analysis of the variogram.

6.3 Spectral Enhancement

Spectral enhancement modified more than one band of data. They can be used to:

- compress bands of data that are similar;
- extract new bands of data that are more interpretable to the eye;
- apply mathematical transforms and algorithms;
- display a wider variety of information in the three available color guns (R, G, B) ERDAS, 2002).

The common techniques apply for spectral enhancement here be explained include (1) Band Ratio, (2) Principal Component Analysis (PCA), (3) Indices and (4) RGB to IHS Transformation and Back Again.

6.3.1 Band Ratioing

Sometime differences in brightness value from identical surface materials are caused by topographic slope and aspect, shadow, or seasonal change in sun illumination angle and intensity. These conditions may hamper ability of interpreter or classification algorithm to identify correctly surface materials or land use in a remotely sensed image. Fortunately, ratio transformation of remotely sensed data can be applied to reduced the effects of such environmental conditions (Jensen, 2005). In addition to minimizing the effects environmental factors, ratios may also provide unique information not available in any single band that is useful discrimination between soils and vegetation.

The mathematical expression of the ratio function is:

$$BV_{i,j,ratio} = \frac{BV_{i,j,k}}{BV_{i,j,l}} \quad (6\text{-}35)$$

where:

$BV_{i,j,k}$ is the original input brightness value in band k

$BV_{i,j,l}$ is the original input brightness value in band l

$BV_{i,j,ratio}$ is the ratio output brightness value

Unfortunately, the computation is not always simple since $BV_{ij} = 0$ is possible. The way to overcome this problem is simply to give any BV_{ij} with a value of 0 the value of 1. To encode the ration value in standard 8-bit format, normalizing functions are applied as follow:

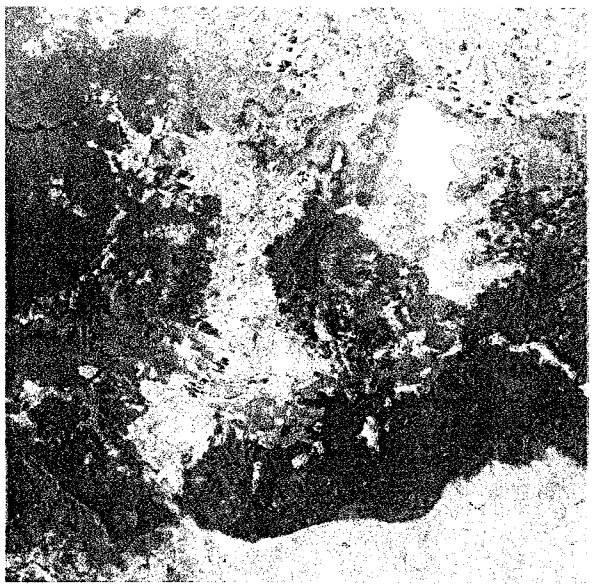
1) Ratio values within the range 1/255 to 1 are assigned values between 1 and 128 by the function:

$$BV_{i,j,n} = Int\left[\left(BV_{i,j,r} \times 127\right) + 1\right] \quad (6\text{-}36)$$

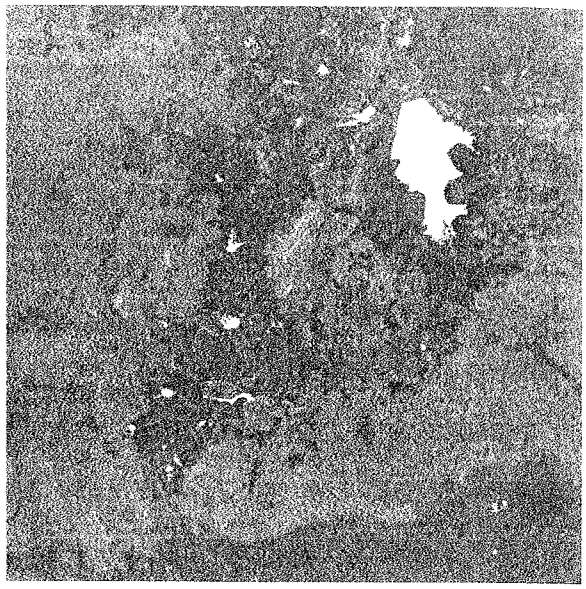
2) Ratio values from 1 to 255 are assigned values within the range 128 to 255 by the function:

$$BV_{i,j,n} = Int\left(128 + \frac{BV_{i,j,r}}{2}\right) \quad (6\text{-}37)$$

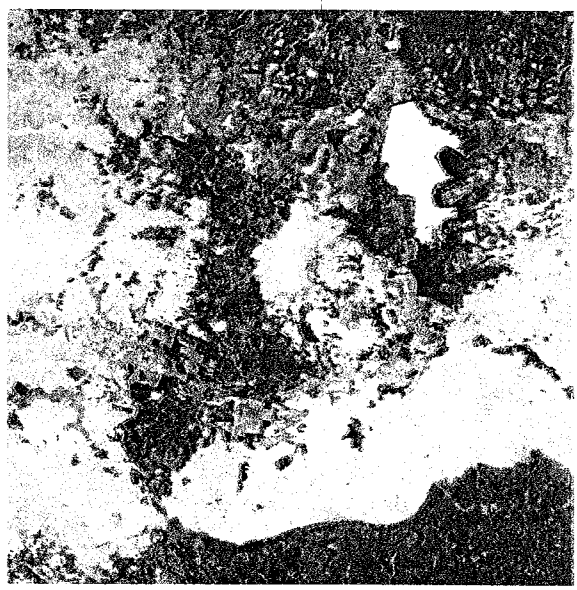
Decide which two bands to ratio is not always a simple task. Often, the analyst simply display various ratios and then select the most visually appealing. The optimum index factor (OIF) and Sheffield Index can be used to identify optimum for band ratio (Jensen, 2005). Example of band ratio was shown in Figure 6-28.



(a) Ratio TM band3/4



(b) Ratio TM band3/5



(c) Ratio TM band4/5

Figure 6-28: The ratio of various Landsat TM bands: Date 6 March 2005, Non Din Deang, Buriram, Thailand (a) Band3/4 (b) Band3/5 and (c) Band4/5

6.3.2 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a technique that transformation of the original remote sensed dataset into a substantially smaller and easier to interpret set of uncorrelated variables that represents most of the information present in the original dataset. The first principal component accounts for the maximum proportion of the variance of the original dataset, and subsequent orthogonal component account for the maximum proportion of the remaining variance. PCA is useful for reducing the dimensionality of dataset. Some researchers applied PCA for data compression and noise reduction of hyperspectral dataset (Jensen, 2005).

To perform PCA we apply a transformation to a *correlated* set of multispectral data. The application of the transformation to the correlated remote sensor data will result in another *uncorrelated* multispectral dataset that has certain ordered variance properties. This transformation is conceptualized by considering the two dimensional distribution of pixel values remote sensor data which labels X_1 and X_2 bands as shown in Figure 6-29.

The spatial relationships between the first two principal components are described as follow:

(a) Scatterplot of data points collected from two remotely bands labeled X_1 and X_2 with the means of the distribution labeled μ_1 and μ_2.

(b) A new coordinate system is created by shifting the axes to an X' system. The values for the new data points are found by the relationship $X_1' = X_1 - \mu_1$ and $X_2' = X_2 - \mu_2$.

(c) The X' axis system is then rotated about its origin (μ_1, μ_2) so that PC_1 is projected through the semi-major axis of the distribution of points and the variance of PC_1 is a maximum. PC_2 must be perpendicular to PC_1. The PC axes are the principal components of this two-dimensional data space.

Component 1 usually accounts for approximately >90% of the variance, and component 2 accounts for 2% to 10%.

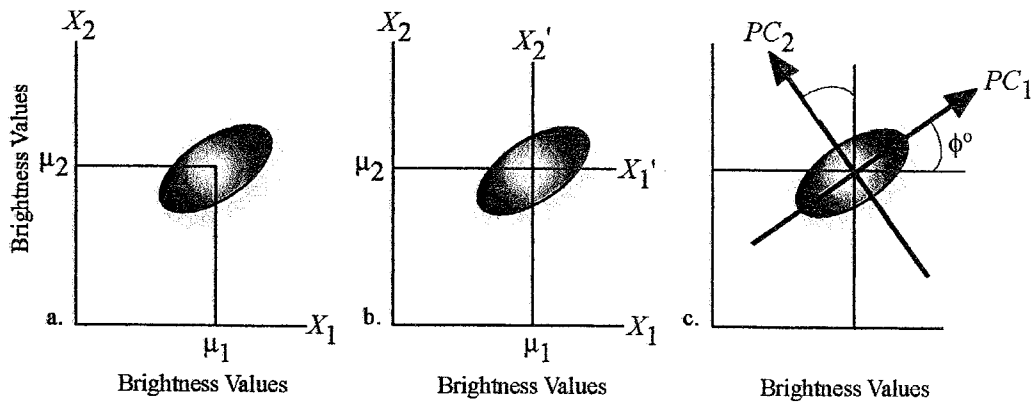


Figure 6-29: Principal Component Analysis (From, Jensen, 2005).

To transform (reproject) the original data on the X_1 and X_2 axes onto the PC_1 and PC_2 axes, we must obtain certain transformation coefficients that we can apply in a linear fashion to the original pixel values. The linear transformation required is derived from the covariance matrix of the original dataset. These coefficients are statistical quantities known as *eigenvectors* or *principal component* (Lillesand et al., 2004).

To compute a principal components transformation, a linear transformation is performed on the data. This means that the coordinates of each pixel in spectral space (the original data file values) are recomputed using a linear equation. The result of the transformation is that the axes in n-dimensional spectral space are shifted and rotated to be relative to the axes of the ellipse. To perform the linear transformation, the eigenvectors and eigenvalues of the n principal components must be mathematically derived from the covariance matrix, as shown in the following equation:

$$E \cdot Cov \cdot ET = V = \begin{bmatrix} v_1 & 0 & 0 & 0 \\ 0 & v_2 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & v_n \end{bmatrix} \quad (6\text{-}38)$$

where:

Cov = the covariance matrix;

E = the matrix of eigenvectors;

T = the transposition function;

V = a diagonal matrix of eigenvalues.

V is computed so that its nonzero elements are ordered from greatest to least, so that $v_1 > v_2 > v_3 \ldots > v_n$.

The matrix V is the covariance matrix of the output principal component file. The zeros represent the covariance between bands (there is none), and the eigenvalues are the variance values for each band. Because the eigenvalues are ordered from v_1 to v_n, the first eigenvalue is the largest and represents the most variance in the data.

Each column of the resulting eigenvector matrix, E, describes a unit-length vector in spectral space, which shows the direction of the principal component (the ellipse axis). The numbers are used as coefficients in the following equation, to transform the original data file values into the principal component values (ERDAS, 2002) as:

$$P_e = \sum_{k=1}^{n} d_k E_{ke} \quad (6\text{-}39)$$

Where:

e = the number of the principal component (first, second),

P_e = the output principal component value for principal component number e,

k = a particular input band,

n = the total number of bands,

d_k = an input data file value in band k,

E_{ke} = the eigenvector matrix element at row k, column e.

Schowengerdt (1997) summarized major characteristics of Principal Component Transform (PCT) as follows:

- It is a rigid rotation in K-D of the original coordinate axes to coincide with the major axes of the data. The data in the PC images result from projecting the original data onto the new axes.
- Although the PC axes are orthogonal to each other in K-D, they are generally not orthogonal when projected to the original multispectral space. This makes it difficult to interpret the eigenvectors of a high dimensional PCT in terms of 2-D scattergrams.
- It optimally redistributes the total image variance in the transformed data. The first PC images contains the maximum possible variance for any linear combination of the original bands, the second PC image contains the maximum possible variance for any axis orthogonal to the first PC, and so forth. This property illustrated in Figure 6-30.
- The transform matrix for PCA is data-dependent. The eigenvalues of a PC calculation for parts of two TM scenes, one largely vegetated and other almost without vegetation are compared in Figure 6-31.

Example of PCA of Landsat-TM data was displayed in Figure 6-32. In this case thermal infrared band 6 was excluded from analysis due to different spatial form others bands.

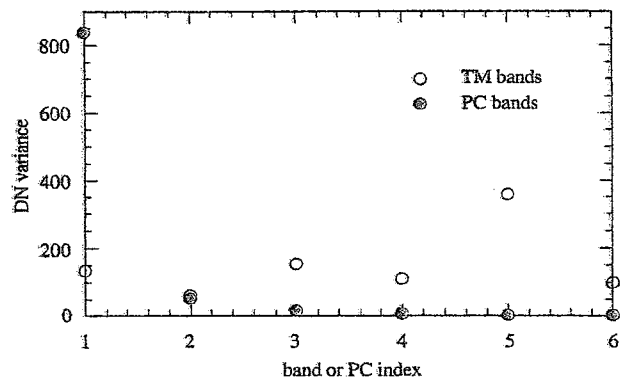


Figure 6-30: Distribution of total image variance across the original spectral bands and across the principal components (From, Schowengerdt, 1997).

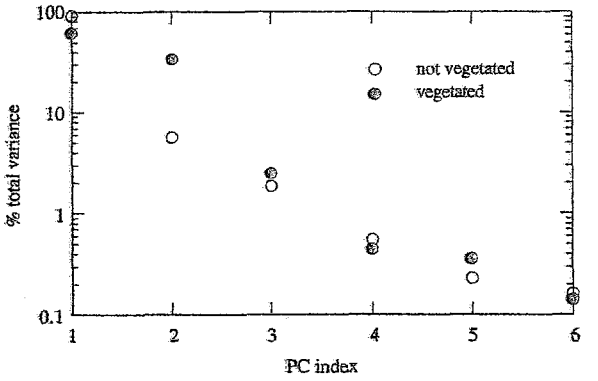


Figure 6-31 Scene dependence in the percentage of total TM image variance captured by each eigenvalue. The TIR band is excluded (From Schwengerdt, 1997).

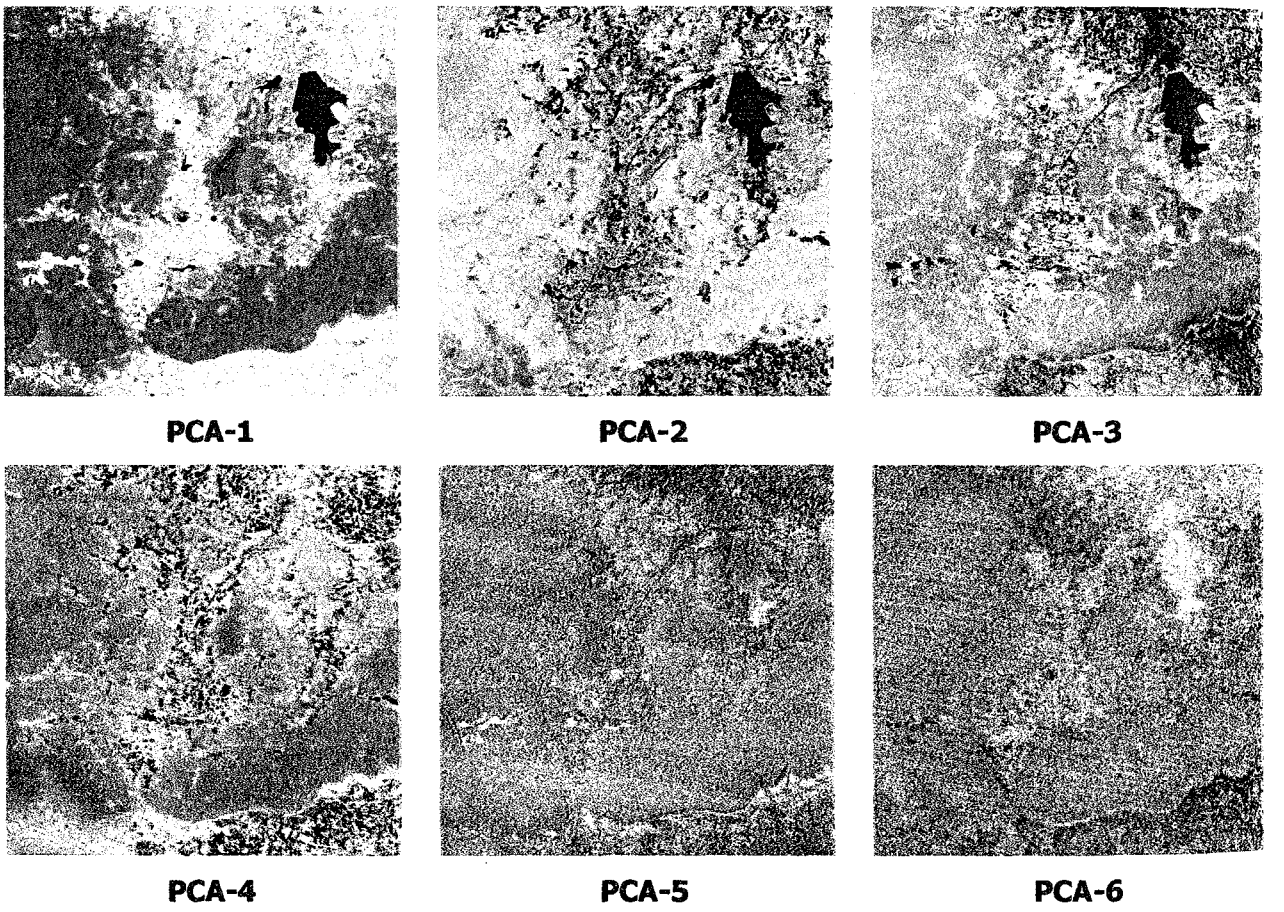


Figure 6-32: Six principal component images of Landsat-TM (TIR Band 6 is excluded): Date 2 Febuary 2005, Non Din Deang, Buriram, Thailand

6.3.3 Indices

Indices are used to create output images by mathematically combining the DN values of different bands. These may be simplistic as (Band X / Band Y) or more complex as [(Band X - Band Y)]/[(Band X + Band Y)]. These ratio images are derived from the absorption/reflection spectra of the material of interest. The absorption is based on the molecular bonds in the (surface) material. Thus, the ratio often gives information on the chemical composition of the target (ERDAS, 2002).

Indices are used extensively in mineral exploration and vegetation analysis to bring out small differences between various rock types and vegetation classes. In many cases, judiciously chosen indices can highlight and enhance differences that cannot be observed in the display of the original color bands.

Indices can also be used to minimize shadow effects in satellite and aircraft multispectral images. Black and white images of individual indices or a color combination of three ratios may be generated.

Certain combinations of TM ratios are routinely used by geologists for interpretation of Landsat TM imagery for mineral type. For example:

- Iron Oxide = TM 3/1
- Clay Minerals = TM 5/7
- Ferrous Minerals = TM 5/4
- Mineral Composite = TM 5/7, 5/4, 3/1
- Hydrothermal Composite = TM 5/7, 3/1, 4/3

Selected remote sensing vegetation indices that compiled by Jensen (2005) were summarized in Table 6-6. Various indices of Landsat TM was demonstrated in Figure 6-33.

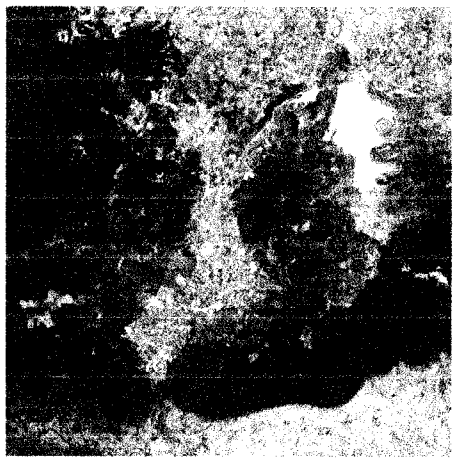
Table 6-6: Selected remote sensing vegetation indices.

Vegetation index	Equation	Note
Simple Ratio (SR)	$SR = \frac{\rho_{red}}{\rho_{nir}}$	ρ_{red} is red reflectance flux ρ_{nir} is NIR reflectance flux
Normalized Differential Vegetation Index (NDVI)	$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}}$	ρ_{red} is red reflectance flux ρ_{nir} is NIR reflectance flux
Kauth-Thomas Tasseled Cap Transformation: MSS data	B = 0.332MSS1 + 0.603MSS2 + 0.675MSS3 + 0.262MSS4 G = -0.283MSS1 - 0.660MSS2 + 0.557MSS3 + 0.338MSS4 Y = -0.899MSS1 + 0.428MSS2 + 0.076MSS3 - 0.041MSS4 N = -0.016MSS1 + 0.131MSS2 - 0.452MSS3 + 0.882MSS4	B is soil brightness index, G is greenness vegetation index, Y is yellow stuff index, N is Non-such.
Kauth-Thomas Tasseled Cap Transformation: TM data	B = 0.2909TM1 + 0.2493TM2 + 0.4806TM3 + 0.5568TM4 + 0.4438TM5 + 0.1706TM7 G = -0.2728TM1 - 0.2174TM2 - 0.5508TM3 + 0.7221TM4 + 0.0733TM5 - 0.1648TM7 W = 0.1446TM1 + 0.1761TM2 + 0.3322TM3 + 0.3396TM4 - 0.6210TM5 – 0.4186TM7	B is brightness, G is greenness, W is wetness.
Infrared Index (II)	$II = \frac{NIR_{TM4} - MIR_{TM5}}{NIR_{TM4} + MIR_{TM5}}$	To study the change in plant biomass and water stress.
Perpendicular Vegetation Index (PVI)	$\sqrt{(0.355MSS4 - 0.149MSS2)^2 + (0.355MSS2 - 0.852MSS4)^2}$	To indicate plant development stage

Table 6-6: (continued)

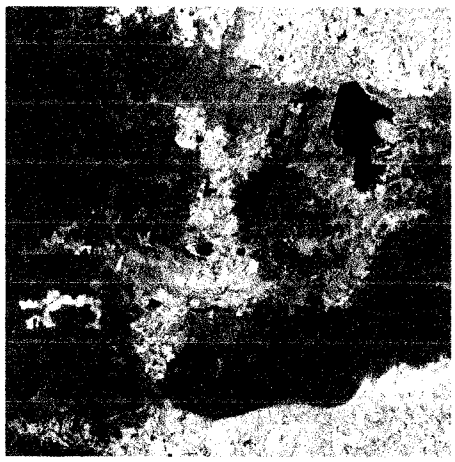
Vegetation index	Equation	Note
Greenness Above Bare Soil (GRABS)	$GRABS = G - 0.09178\ B + 5.58959$	G is Kauth-Thomas transform greenness B is Kauth-Thomas transform soil brightness
Moisture Stress Index (MSI)	$MSI = \frac{MidIR_{TM5}}{NIR_{TM4}}$	ρ_{red} is red reflectance flux ρ_{nir} is NIR reflectance flux
Leaf Water Content Index (LWCI)	$LWCI = \frac{-\log\left[1-(NIR_{TM4}-MidIR_{TM5})\right]}{-\log\left[1-(NIR_{TM4\,ft}-MidIR_{TM5\,ft})\right]}$	where ft represent reflectance in the specified bands when leaves are at their maximum relative water content (RWC). RWC defined as: $RWC = \frac{field\_weight - oven\_dry\_weight}{turgid\_weight - oven\_dry\_weight} \times 100$
Soil Adjusted Vegetation Index (SAVI)	$SAVI = \frac{(1+L)(\rho_{nir}-\rho_{red})}{\rho_{nir}+\rho_{red}+L}$	L is a canopy background adjustment factor
Soil and Atmospherically Resistant Vegetation Index (SARVI)	$SARVI = \frac{p*nir - p*rb}{p*nir + p*rb}$	Where: $p*rb = p*red - \gamma(p*blue - p*red)$
Enhanced Vegetation Index (EVI) : MODIS	$EVI = G\frac{(\rho^*_{nir}-\rho^*_{red})}{\rho^*_{nir}+C_1\rho^*_{red}-C_2\rho^*_{blue}+L}(1+L)$	G is greenness L is soil adjustment factor (1.0) C_1 is 6.0 and C_2 is 7.5

Modified from Jensen (2005)

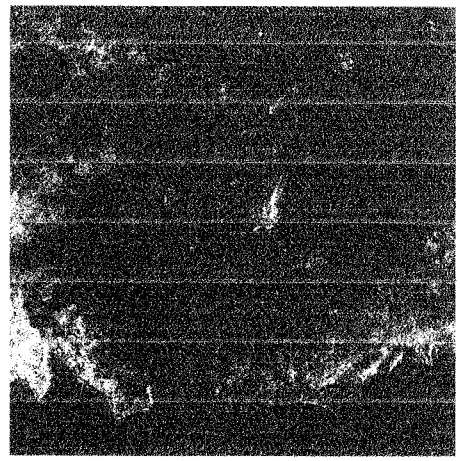


(a) Simple Ratio (SR)

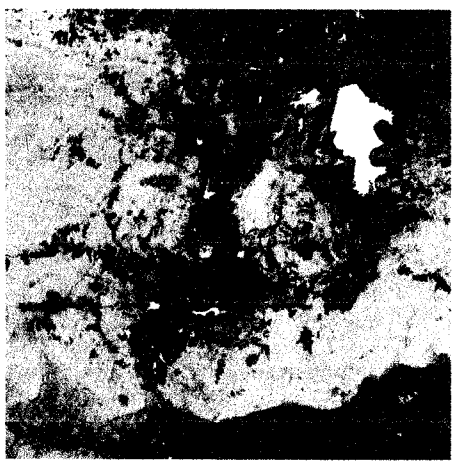
(b) Normalized Differential Vegetation Index (NDVI)



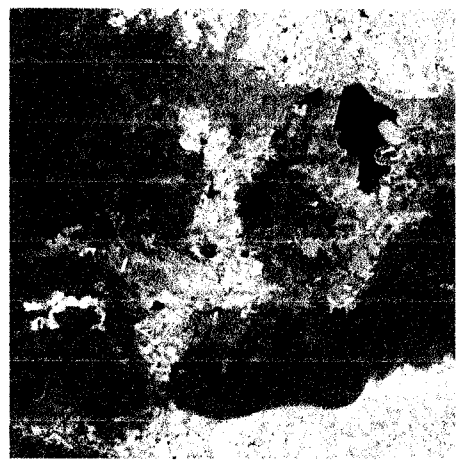
(c) Kauth-Thomas Tasseled Cap: Brightness



(d) Kauth-Thomas Tasseled Cap: Greenness



(e) Kauth-Thomas Tasseled Cap: Wetness



(f) Greenness Above Bare Soil (GRABS)

Figure 6-33: Various vegetation indices of Landsat TM: Date 2 Febuary 2005, Non Din Deang, Buriram, Thailand (a) SR (b) NDVI (c-e) Kauth-Thomas Tasseled Cap and (f) GRABS

6.3.4 RGB to IHS Transformation and Back Again

The intensity-hue-saturation color coordinate system is based on a hypothetical color sphere as shown in Figure 6-34. The vertical axis represents **intensity (I)** which varies from black (0) to white (255) and is not associated with any color. The circumference of the sphere represents **hue (H),** which is the dominant wavelength of color. Hue values begin with 0 at the midpoint of red tones and increase counterclockwise around the circumference of the sphere to conclude with 255 adjacent to 0. **Saturation (S)** represents the **purity** of the color and ranges from 0 at the center of the color sphere to 255 at the circumference. A saturation of 0 represents a completely impure color in which all wavelengths are equally represented and which the eye will perceive as a shade of gray that ranges from white to black depending on intensity. The relationship between the RGB and HIS systems is shown in Figure 6-35. IHS values can be derived from the RGB values through the transformation equations (Sabins, 1987):

$$I = R + G + B \tag{6-40}$$

$$H = \frac{G - B}{I - 3B} \tag{6-41}$$

$$S = \frac{I - 3B}{I} \tag{6-42}$$

Jensen (2005) stated that the IHS transformation is often used to merge multiple types of remote sensor data. The method generally involves four steps:

1. RGB to HIS. Three bands of lower-spatial-resolution remote sensor data in RGB color space are transformed into three bands in HIS color space.

2. Contrast manipulation. The high-spatial-resolution image is contrast stretched so that it has approximately the same variance and mean as the intensity (I) image.

3. Substitution. The stretched, high-spatial-resolution image is substituted for the intensity (I) image.

4. HIS to RGB. The modified HIS dataset is transformed back into RGB color space using an inverse HIS transformation. The justification for replacing the intensity (I) component with the stretched higher-spatial-resolution images have approximately the same spectral characteristics as shown in Figure 6-36.

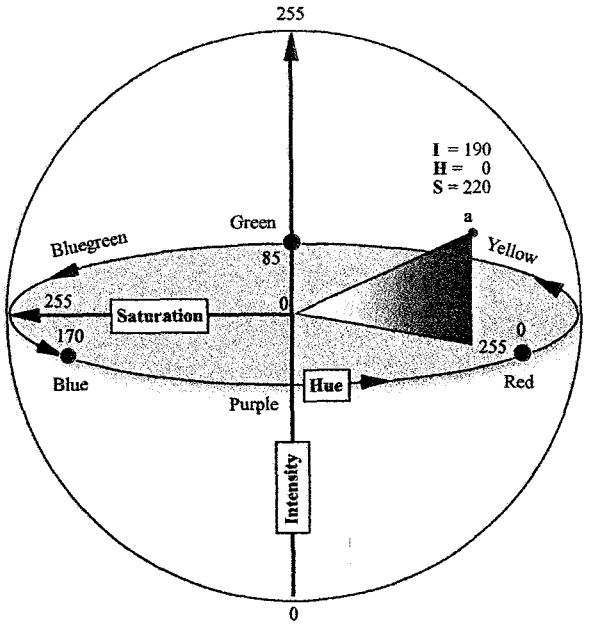


Figure 6-34: Intensity-hue-saturation (HIS) color coordinate system (From Jensen, 2005)

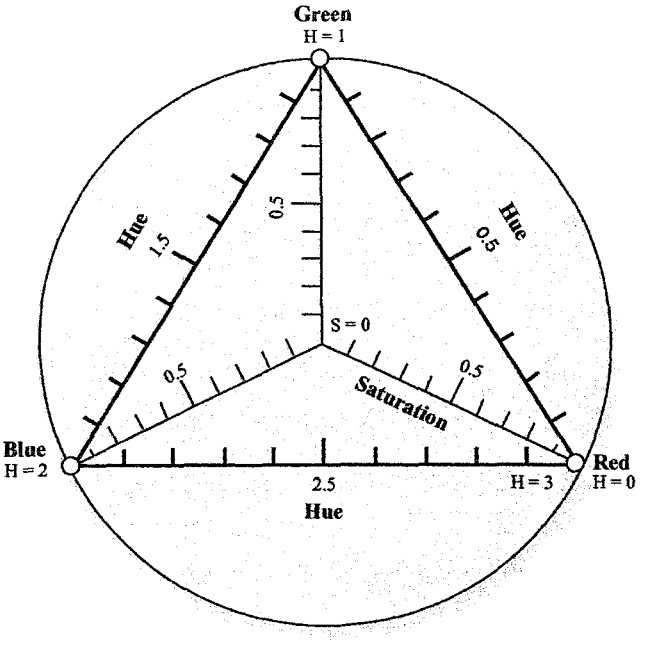


Figure 6-35: Relationship between intensity-hue-saturation (HIS) color coordinate system and RGB coordinate system (From Jensen, 2005)

Figure 6-36: Resolution merge between SPOT data (25 Jan 2006) and Landsat TM (9 Mar 2006) of Non Din Deang, Buriram, Thailand based on modified IHS resolution merge.

Chapter 7: Image Classification in Digital Image Processing

7.1 Introduction

Classification is the process of partitioning an image data set into a discrete number of class accordance with specific criteria that are based, in part, on the individual image point data value. The image data set is the set of data values for all of the pixels in the image and the image pixel data values is the set of data values associated with an individual pixel. Classification is the process of using the pixel data values to allocate that pixel to one of a discrete number of classes.

McCloy (2006) stated that classification contained a number of important concepts, the underlying assumption and philosophies of which need to be understood to achieved reliable and consistent results. The concepts are:

(1) The use of numerical models is used to establish decision surfaces. Partitioning the data sets into class domains is usually done by means of classifiers algorithm. The shape and position of the decision surfaces depend on the characteristics of the algorithm. It is essential to understand the theoretical model on which the classifier algorithm is based, and be sure that the actual data is a satisfactory fit to this model.

(2) Use of parameters for those models that are usually derived from field data. The values used for the algorithm parameters significantly affect the shape and position of the decision surfaces. Estimation of parameter values is usually done by use of a sample of data selected for that purpose, or training data. The sample of data must be representative of the class, that is, it needs to represent the range of conditions that apply to the class, and it must be of an adequate size to give representative value for the class. An important factor to consider in selecting the size of the sample is the sensitivity of the algorithm to slight variations in the derived parameter values.

(3) Decisions are based on "Spectral" data. Classifier algorithms allocate a pixel to a class either on the data set belonging to the pixel, or they can utilize additional data about the pixel or data derived from the surrounding pixels. Most classifiers limit classification to the use of the spectral data for the pixel and as such are called spectral classifiers. However, there are many situations where other data sets can significantly improve classification accuracy. For example, certain classes of tree species may have similar spectral values in the image data, but they may grow in distinct climatic regimes reflected in differences in elevation, gradient and aspect or temperature and rainfall, or a combination of these. In a situation like this, inclusion of additional data sets, effectively as another band of data, can improve the classification accuracy.

(4) Partitioning spectral domain into discrete classes. Once the classification parameters are defined they create decision surfaces within the data domain such that all pixel value within the one sub-domain, as specified by the decision surface, will be assigned to the one class. The resulting classification is uniform within the class sub-domains. In some circumstances this partitioning is a valid reflection of actual conditions. In other situations, when the different conditions are not so discrete, then partitioning the image in this way introduces artificialities that can introduce errors in the classification. Variations are a gradual change from one to another as part of a continuum of conditions within the one class such pasture class. Many of spectral classifiers cannot handle these within-class variations in a satisfactory way for a number of reasons include (a) inadequate training data (b) poor fit of the data to the normal distribution, (c) Inadequacy of spectral classes and (d) the existence of mixed pixels.

7.2 Classification Process

Thematic classification of an image involved several steps included three major components: feature extraction, training and labeling (Schowengerdt, 1997) (Figure 7-1).

Feature extraction: Transform the multiple images by a spatial image by a spatial or spectral transform to a feature image. Examples are selection of a subset of bands, a Principal Component Transform to reduce the data dimensionality, or a spatial smoothing filter. This step is optional, i.e. the multispectral image can be used directly, if desired.

Training: Extract the pixels to be used for training the classifier to recognize certain categories, or classes. Determine the discriminant functions in the feature space which allow assignment of a class label to each pixel. This step is either supervised by the analyst or unsupervised.

Labeling: Apply the discriminant functions to the entire feature image and label all pixels. If the training was supervised, the labels are associated with the discriminant functions; if it was unsupervised, the analyst must now supervise the labeling. The output map consists of one label for each pixel.

The end result is to convert the numerical image data into descriptive labels that categorize different surface materials or conditions. A large reduction in data quantity takes places. The map can therefore be stored with binary encoding as a single band file with less than 8 bits/pixel. By virtue of labeling process, however, we have presumably converted the data into something that has informational value.

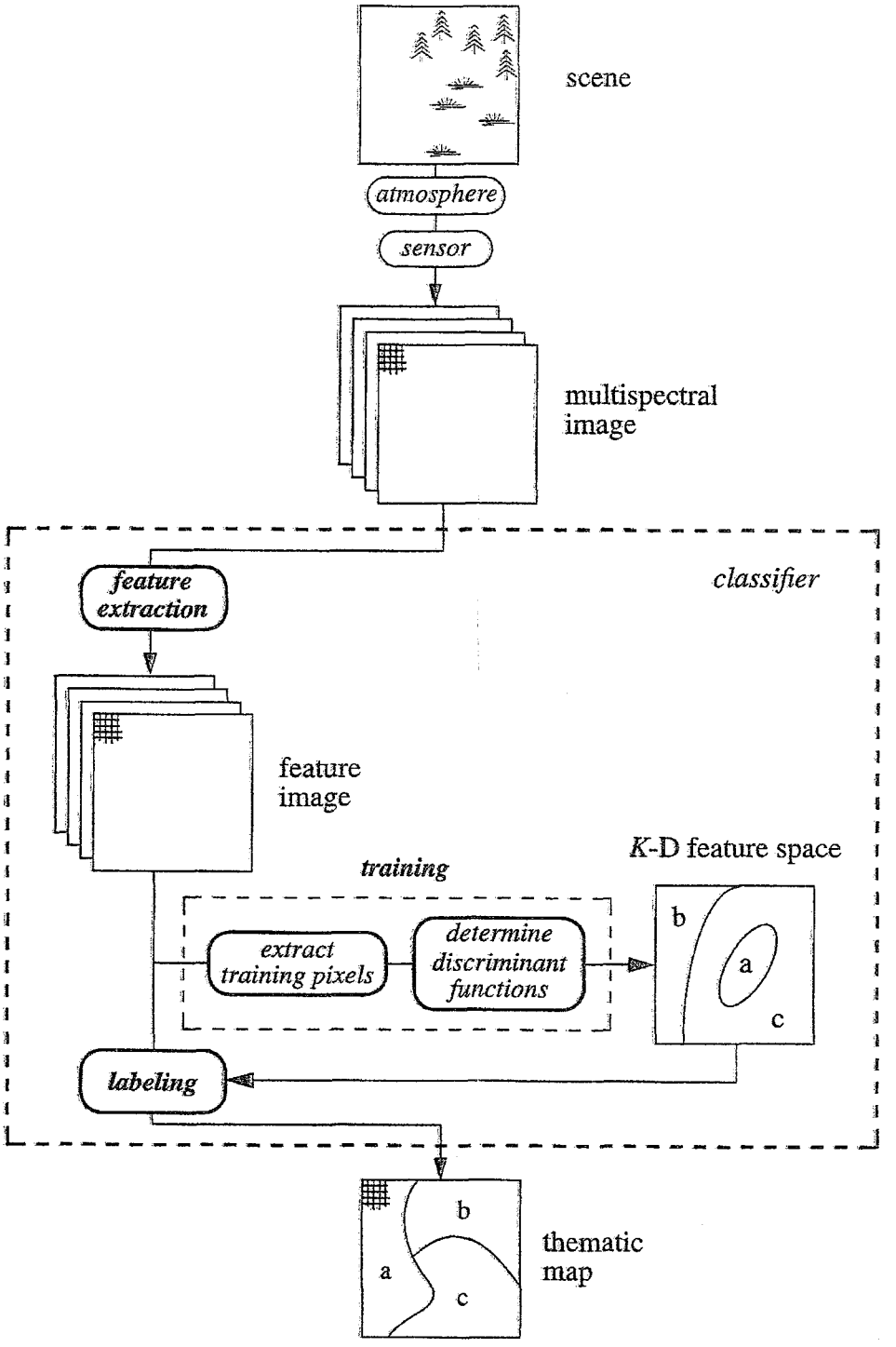


Figure 7-1: The data flow in a classification process (from Schowengerdt, 1997)

Jensen (2005) suggested the general steps required to extract land-cover information from digital multispectral remote sensor data as following:

1. **State the nature of the classification problem.**
    - Specify the geographic region of interest.
    - Define the classes of interest.
    - Determine if it is to be a hard or fuzzy classification.
    - Determine if it is to be a per-pixel or object-oriented classification.
2. **Acquire appropriate remote sensing and initial ground reference data.**
    - Select remotely sensed data based on the following criteria
        - Remote sensing system considerations
            - Spatial, spectral, temporal, and radiometric resolution
        - Environmental considerations
            - Atmospheric, soil moisture, phenological cycle, etc.
    - Obtain initial ground reference data based on:
        - A priori knowledge of the study area
3. **Process remote sensor data to extract thematic information;**
    - Radiometric correction or normalization
    - Geometric correction
    - Select appropriate image classification logic
        - Parametric methods
        - Nonparametric methods
        - Nonmetric methods
    - Select appropriate image classification algorithm
        - Supervised classification
        - Unsupervised classification
        - Hybrid classification
    - Extract data from initial training site (if required)
    - Select most appropriate bands using feature selection criteria
        - Graphical
        - Statistical
    - Extract training statistics and rules based
        - Final band selection (if required), and/or
        - Machine-learning
    - Extract thematic information

- For each pixel or for each segmented image object (supervised)
- Label pixels or image object (unsupervised)

4. **Perform accuracy assessment;**
   - Select method:
     - Qualitative confidence-building
     - Statistical measurement
   - Determine number of samples required by class
   - Select sampling scheme
   - Obtain ground reference test information
   - Create and analyze error matrix
     - Univariate and multivariate statistical analysis
5. **Accept or reject previous stated hypothesis;**
6. **Distributed results if accuracy is acceptable.**

Jensen (2005) claimed that the actually multispectral classification may be performed using a variety methods, including

- Algorithms based on parametric and nonparametric statistics that use ratio-and interval scaled data and Nonmetric methods that can also incorporate nominal scale data,
- The use of supervised or unsupervised classification logic
- The use of hard or soft (fuzzy) set classification logic to create hard or fuzzy thematic output products,
- The use of per-pixel or object-oriented classification logic, and
- Hybrid approaches.

Description of the above mentioned methods are here briefly summarized in following sections based on Jensen (2005).

Parametric methods: Parametric methods such as maximum likelihood classification and unsupervised clustering assume normally distributed remote sensor data and knowledge about the forms of the underlying class density functions.

Nonparametric methods: Nonparametric methods such as nearest-neighbor classifiers, fuzzy classifiers, and neural networks may be applied to remote sensor data that are not normally distributed and without the assumption that the forms of the underlying densities are known.

Nonmetric methods: Nonmetric methods such as rule-based decision tree classifiers can operate on both real-valued data (e.g., reflectance values from 0 to 100%) and nominal scaled data (e.g., class 1 = forest; class 2 = agriculture).

Supervised classification algorithm: In a supervised classification, the identity and location of the land-cover types are known a priori through a combination of fieldwork, interpretation of aerial photography, map analysis, and personal experience. The analyst attempts to locate specific sites in the remotely sensed data that represent homogeneous examples of these known land-cover types. These areas are commonly referred to as training sites because the spectral characteristics of these known areas are used to train the classification algorithm for eventual land-cover mapping of the remainder of the image. Multivariate statistical parameters (means, standard deviations, covariance matrices, correlation matrices, etc.) are calculated for each training site. Every pixel both within and outside the training sites is then evaluated and assigned to the class of which it has the highest likelihood of being a member.

Unsupervised classification algorithm: In an unsupervised classification, the identities of land-cover types to be specified as classes within a scene are not generally known a priori because ground reference information is lacking or surface features within the scene are not well defined. The computer is required to group pixels with similar spectral characteristics into unique clusters according to some statistically determined criteria. The analyst then re-labels and combines the spectral clusters into information classes.

Also, the analyst should considered and selected appropriate classification logics based on hard or soft classification logic.

Hard classification logic: Supervised and unsupervised classification algorithms typically use hard classification logic to produce a classification map that consists of hard, discrete categories (e.g., forest, agriculture).

Fuzzy classification logic: Fuzzy classification produces thematic output products that contain fuzzy information. Fuzzy classification is based on the fact that remote sensing detectors record the reflected or emitted radiant flux from heterogeneous mixtures of biophysical materials such as soil, water, and vegetation found within IFOV. The land cover classes found within the IFOV (pixel) often grade into one another without sharp, hard boundaries. Instead of being assigned to just a single class out of m possible classes, each pixel in a fuzzy classification has m membership grade values that describe the proportion of the m land cover types found within the pixel (e.g., 10% bare soil, 10% scrub-shrub, 80% forest). Comparison of hard and soft classification logics is shown in Figure 7-2.

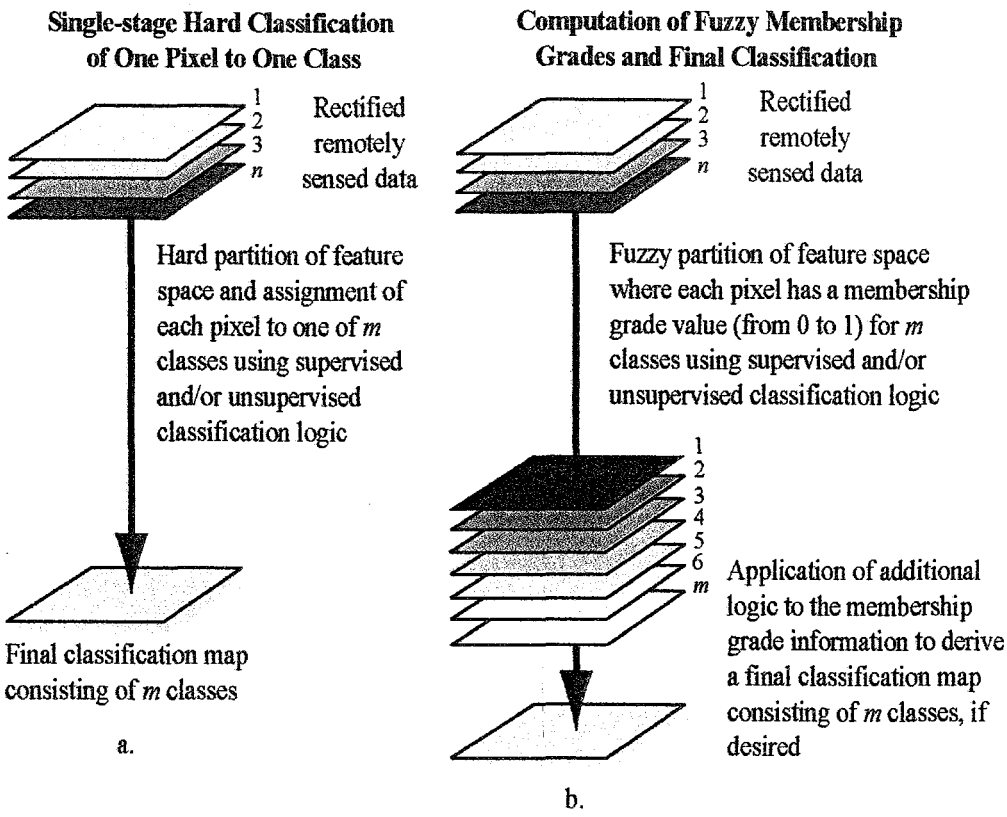


Figure 7-2: Comparison between a traditional hard and soft classification logic (From Jensen, 2005)

Furthermore, in the past, most digital image classification was based on processing the entire scene pixel by pixel. This is commonly referred to as **per-pixel classification**. However, **Object-oriented classification** techniques allow the analyst to decompose the scene into many relatively homogenous image objects (referred to as patches or segments) using a multi-resolution image segmentation process. The various statistical characteristics of these homogeneous image objects in the scene are then subjected to traditional statistical or fuzzy logic classification. Object-oriented classification based on image segmentation is often used for the analysis of very high-spatial-resolution imagery (Jensen, 2005).

7.3 Land Use and Land Cover Classification Scheme

Land use and land cover classification based on statistical pattern recognition techniques applied to multispectral remotely sensed data is one of the most often used methods of information extraction. Land use and land cover information are often used by scientists and researches in various studies. Land cover refers to the type of material present on the landscape (e.g., water, sand, crops, forest, wetland, human-made materials such as asphalt). Land use refers to what people do on the land surface (e.g., agriculture, commerce, settlement). The pace, magnitude, and scale of human alterations of the Earth's land surface are unprecedented in human history.

Therefore, land-cover and land-use data are central to such United Nations' Agenda 21 issues as combating deforestation, managing sustainable settlement growth, and protecting the quality and supply of water resources (Jensen, 2005).

Jensen (2005) stated that all classes of interest must be selected and defined carefully to classify remotely sensed data successfully into land-use and/or land-cover information. This requires the use of a classification scheme containing taxonomically correct definitions of classes of information that are organized according to logical criteria. If a hard classification is to be performed, then the classes in the classification system should normally be:

- mutually exclusive,
- exhaustive, and
- hierarchical.

Mutually exclusive means that there is no taxonomic overlap (or fuzziness) of any classes (i.e., deciduous forest and evergreen forest are distinct classes). Exhaustive means that all land-cover classes present in the landscape are accounted for and none have been omitted. Hierarchical means that sublevel classes (e.g., single-family residential, multiple-family residential) may be hierarchically combined into a higher- level category (e.g., residential) that makes sense. This allows simplified thematic maps to be produced when required.

It is also important for the analyst to realize that there is a fundamental difference between information classes and spectral classes. Information classes are those that human beings define. Conversely, Spectral classes are those that are inherent in the remote sensor data and must be identified and then labeled by the analyst (Jensen,2005).

Certain land use and land cover classification have been developed can be applied for interpreting and analyzing remotely sensed data such as USGS Land-Use/Land-Cover Classification System for Use with Remote Sensor Data (Table 7-1). Jensen (2005) suggested that if a reputable classification system already exists, it is foolish to develop an entirely new system that will probably be used only by ourselves. It is better to adopt or modify existing nationally or internationally recognized classification systems. This allows us to interpret the significance of our classification results in light of other studies and makes it easier to share data. In addition, it should be reiterated that there is a general relationship between the level of detail in a classification scheme and the spatial resolution of remote sensor systems used to provide information. Basically, the level of detail in desired classification system dictated the spatial resolution of remote sensor data that should be used as shown in Figure 7-3.

Table 7-1: USGS Land-Use/Land-Cover Classification System for Use with Remote Sensor Data

Level I	Level II
1 Urban or built-up land	11 Residential
	12 Commercial and service
	13 Industrial
	14 Transportation, communications, and utilities
	15 Industrial and commercial complexes
	16 Mixed urban or built-up land
	17 Other urban or built-up land
2 Agricultural land	21 Cropland and pasture
	22 Orchards, groves, vineyards, nurseries, and ornamental horticultural areas
	23 Confined feeding operations
	24 Other agricultural land
3 Rangeland	31 Herbaceous rangeland
	32 Shrub and brush rangeland
	33 Mixed rangeland
4 Forest land	41 Deciduous forest land
	42 Evergreen forest land
	43 Mixed forest land
5 Water	51 Streams and canals
	52 Lakes
	53 Reservoirs
	54 Bays and estuaries
6 Wetland	61 Forested wetland
	62 Non-forested wetland
7 Barren land	71 Dry salt flats
	72 Beaches
	73 Sandy areas other than beaches
	74 Bare exposed rock
	75 Strip mines, quarries, and gravel pits
	76 Transitional areas
	77 Mixed barren land
8 Tundra	81 Shrub and brush tundra
	82 Herbaceous tundra
	83 Bare ground tundra
	84 Wet tundra
	85 Mixed tundra
9 Perennial snow or ice	91 Perennial snowfield
	92 Glaciers

From Lillesand et al (2004).

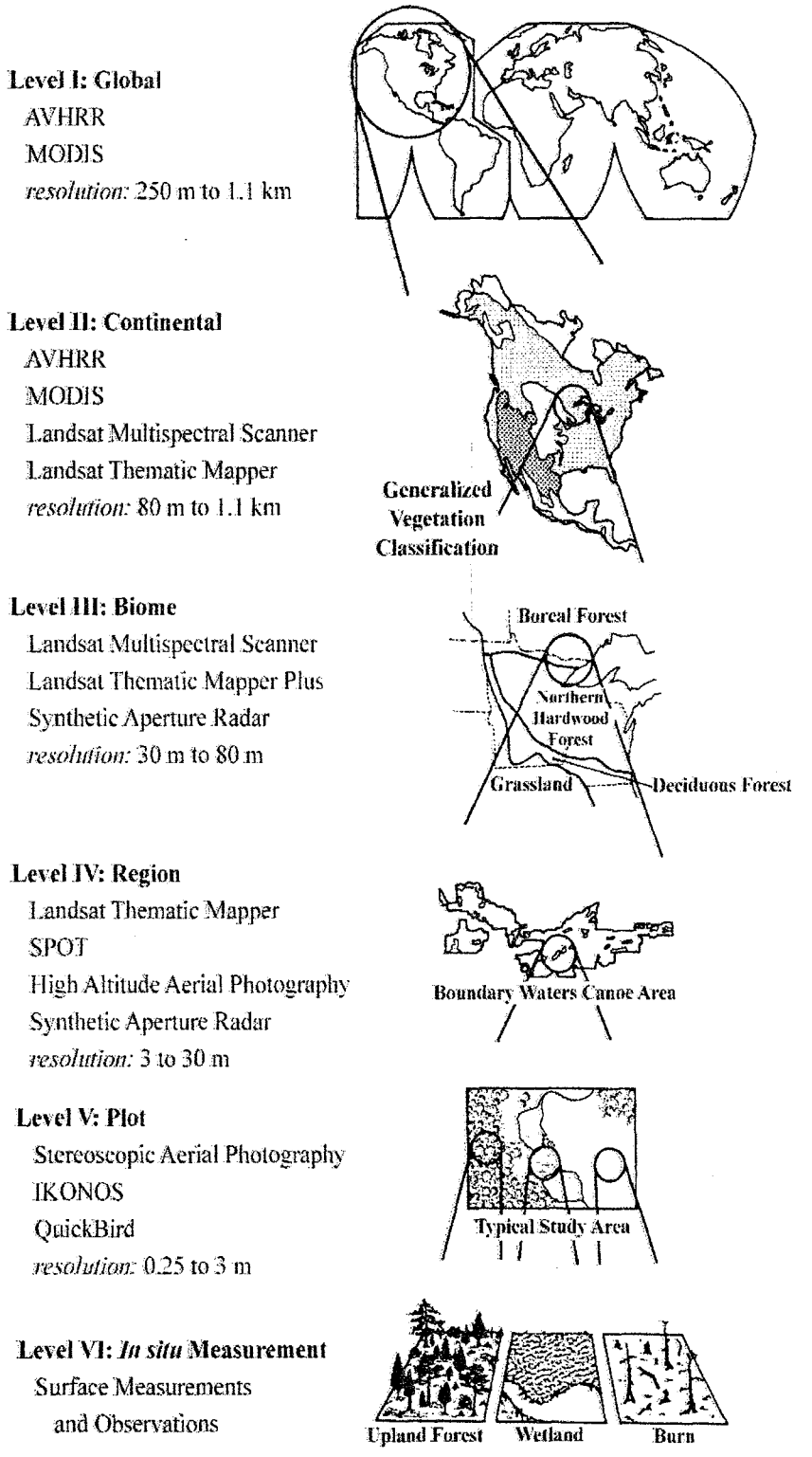


Figure 7-3: Relationship between the level of detail required and the spatial resolution of representative remote sensing system for vegetation inventories (From, Jensen, 2005).

7.4 Training Site Selection and Feature selection

7.4.1 Training Site Selection

ERDAS (2002) suggested that in supervised training, it is important to have a set of desired classes in mind, and then create the appropriate signatures from the data. Analysts must also have some way of recognizing pixels that represent the classes that you want to extract. Supervised classification is usually appropriate when analysts want to identify relatively few classes, when analysts have selected training sites that can be verified with ground truth data, or when analysts can identify distinct, homogeneous regions that represent each class. On the other hand, if analysts want the classes to be determined by spectral distinctions that are inherent in the data so that analysts can define the classes later, then the application is better suited to unsupervised training. Unsupervised training enables analysts to define many classes easily, and identify classes that are not in contiguous, easily recognized regions.

Jensen (2005) also suggested that an analyst may select training sites within the image that are representative of the land-cover or land-use classes of interest after the classification scheme is adopted. The training data should be of value if the environment from which they were obtained is relatively homogeneous. For example, if all the soils in a grassland region are composed of well-drained sandy-loam soil, then it is likely that grassland training data collected throughout the region would be representative. However, if the soil conditions change across the study area, it is likely that grassland training data acquired in the dry-soil part of the study area will not be representative of the spectral conditions for grassland found in the moist-soil portion of the study area. This is called a geographic signature extension problem, meaning that it may not be possible to extend the grassland remote sensing training data through x, y space.

The easiest way to remedy this situation is to apply geographical stratification during the early stages of a project. All significant environmental factors that contribute to geographic signature extension problems should be identified, such as differences in soil type, water turbidity, crop species, unusual soil moisture conditions, atmospheric haze, and so on. Such environmental conditions should be carefully annotated on the imagery and the selection of training sites made based on the geographic stratification of these data. It is necessary to train the classifier over relatively short geographic distances. Each individual stratum may have to be classified separately. The final classification map of the entire region will then be a composite of the individual stratum classifications. However, if environmental conditions are homogeneous or can be held constant, it may be possible to extend signatures vast distances in space, significantly reducing the training cost and effort.

Once spatial and temporal signature extension factors have been considered, the analyst selects representative training sites for each class and collects the spectral statistics for each pixel found within each training site. Each site is usually composed of many pixels. The general rule is that if training data are being extracted from n bands then >10n pixels of training data are collected for each class. This is sufficient to compute the variance–covariance matrices required by some classification algorithms.

There are a number of ways to collect the training site data, including:

- collection of in situ information such as tree type, height, percent canopy closure, and diameter-at-breast-height (dbh) measurements,
- on-screen selection of polygonal training data, and/or
- on-screen seeding of training data.

7.4.2 Feature selection

Once the training statistics have been systematically collected from each band for each class of interest, a judgment must be made to determine the bands that are most effective in discriminating each class from all others. This process is commonly called feature selection. The goal is to delete from the analysis the bands that provide redundant spectral information. In this way the dimensionality in the dataset may be reduced. This minimizes the cost of the digital image classification process (Jensen, 2005). Feature selection may involve both graphical and statistical analysis to determine the degree of between-class separability in training data.

7.4.2.1 Graphic Methods of Feature Selection

Graphical methods to determine the degree of between-class separability in training data were here summarized based on Jensen (2005).

A. Bar graph spectral plots: These methods were one of the first simple feature selection aids where the mean ± standard deviation is displayed in a bar graph format for each band. This provides an effective visual presentation of the degree of between-class separability for one band at a time (Figure 7-4). However, the display provides no information on how well any two bands would perform.

B. Cospectral mean vector plots: Cospectral mean vector plots may be used to present statistical information about at least two bands at one time. Theoretically, the greatest the distance is between numbers in the feature space distribution, the greatest the potential for accurate between-class discrimination. Using this method, only two bands of data can be analyzed at one time. Therefore, they devised an alternative method whereby the size of the numeral depicts the location of the information in a third dimension of feature space. Example of two dimensional feature space plots show in Figure 7-5.

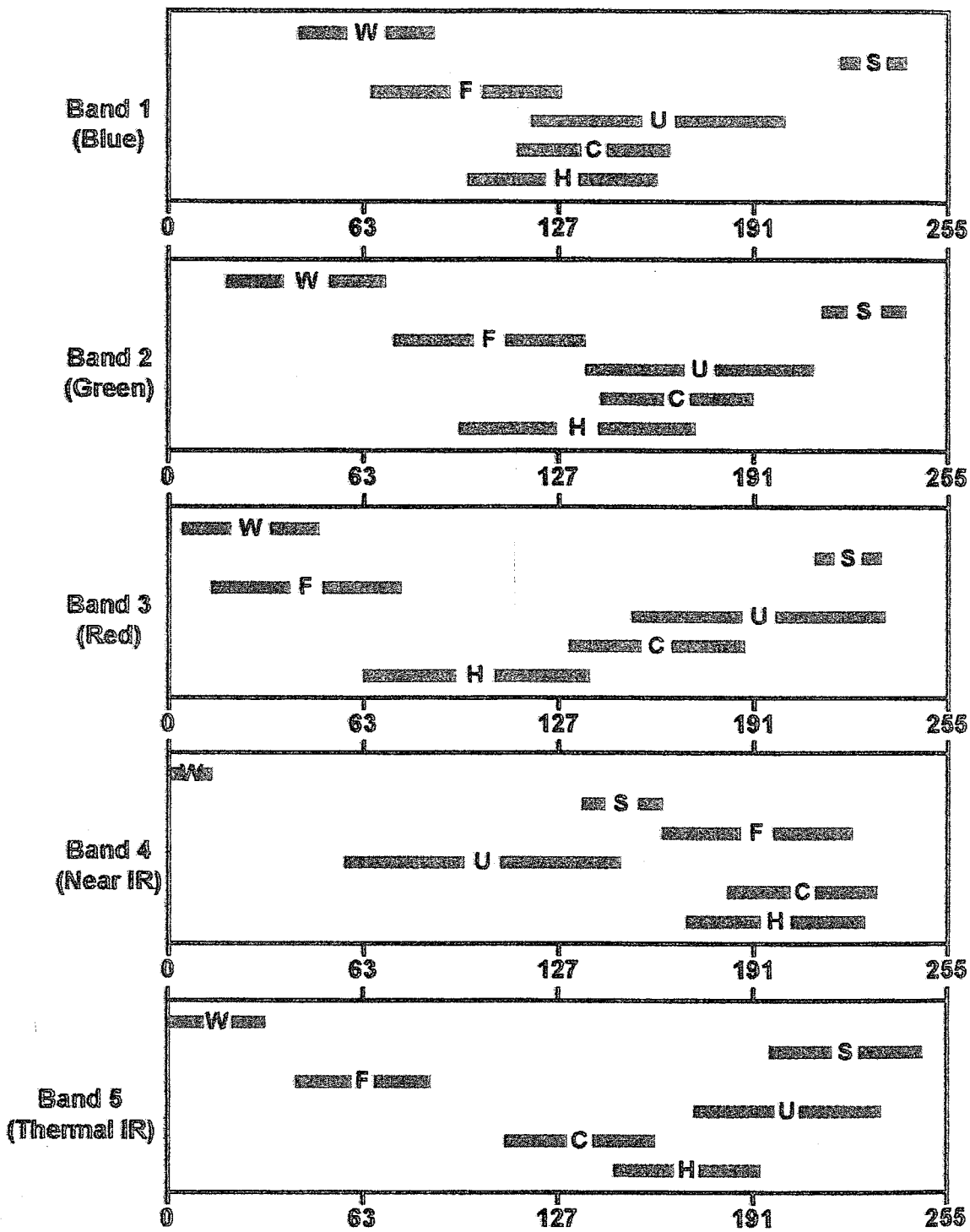


Figure 7-4: Coincident spectral plots for training data obtained in five bands for six cover type (Lillesand et al., 2004)

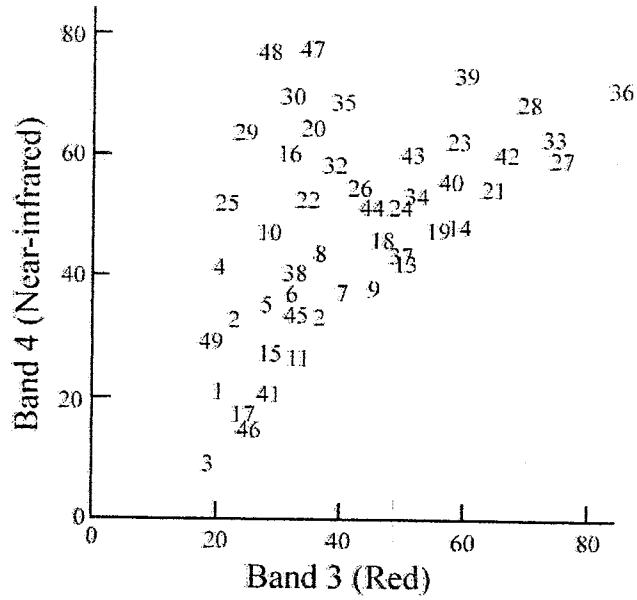


Figure 7-5: Cospectral mean vector plot of 49 clusters (From Jensen, 2005)

C. Cospectral parallelepipeds or ellipse plots: Cospectral parallelepipeds or ellipse plots in two-dimensional feature space provide useful visual between-class separability information. They are created using the mean, μ_{ck}, and standard deviation, σ_{ck}, of training class statistics for each class c and band k. The lower and upper limits of the two-dimensional parallelepipeds (rectangles) were obtained using the mean ± 1 standard deviation of each band for each class. Figure 7-6 shows ellipse plot for signatures evaluation.

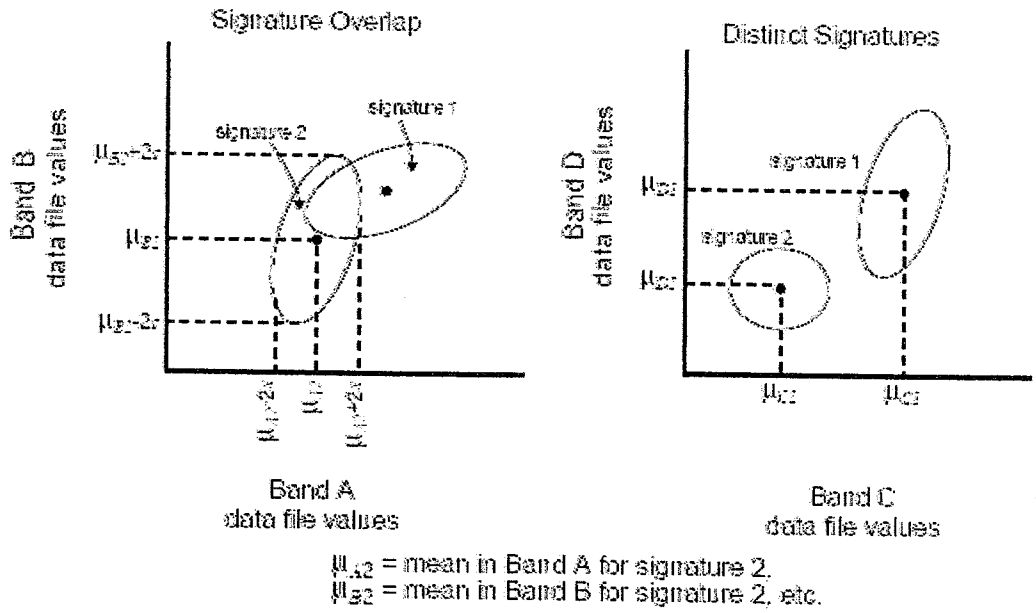


Figure 7-6: Ellipse Evaluation of Signatures (From ERDAS, 2002)

It is possible to display three bands of training data at once using tri-spectral parallelepipeds or ellipse plots in synthetic three-dimensional feature space. Jensen and Toll (1982) present a method of displaying parallelepipeds in synthetic three-dimensional space and of interactively carrying the viewpoint azimuth and elevation angles to enhance feature analysis and selection. Again, the mean, μck, and standard deviation, σck, of training class statistics for each class c and band k are used to identify the lower and upper threshold value for each class and band as shown in Figure 7-7.

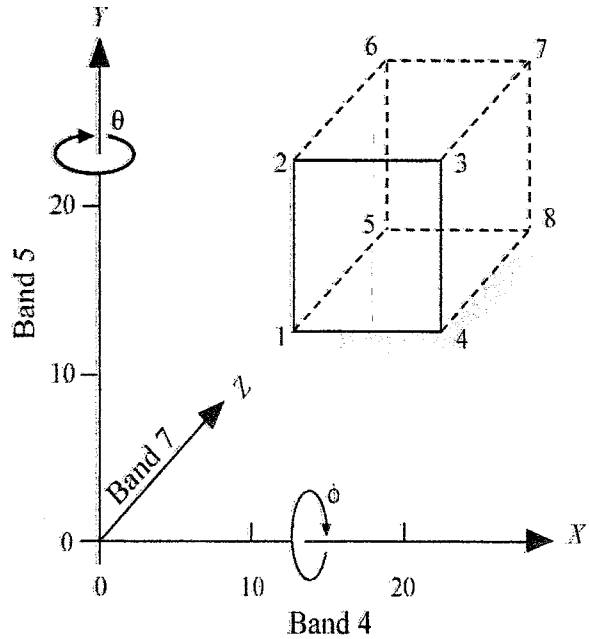


Figure 7-7: Simple parallelepiped displayed in pseudo three-dimensional space (From Jensen, 2005).

7.4.2.2 Statistical Measures of Feature Selection

Statistical methods of feature selection are used to quantitatively select which subset of bands or features provides the greatest degree of statistical separability between any two classes (Jensen, 2005). Schowengerdt (1997) summarized common measure of statistical separability used in digital image processing included (a) city block, (b) Euclidean (c) angular (d) normalized city block (f) Mahalanobis (g) divergence (h) transformed divergence (i) Bhattacharry and (j) Jeffries Matusita as shown in Table 7-2. Example of transformed divergence was shown in Table 7-3.

Table 7-2: Various distance measure methods for separability analysis.

Name	Formula	Remark
City block	$L_1 = \left\|\mu_a - \mu_b\right\| = \sum_{k=1}^{K} \left\|m_{ak} - m_{bk}\right\|$	Simple measures of separation of the means without the covariance.
Euclidean	$L_2 = \left\|\mu_a - \mu_b\right\| = \left[\left(\mu_a - \mu_b\right)^T \left(\mu_a - \mu_b\right)^{1/2}\right]$ $L_2 = \left[\sum_{k=1}^{K} \left(m_{ak} - m_{bk}\right)^2\right]^{1/2}$	
Angular	$ANG = a\cos\left(\frac{\mu_a{}^T \mu_b}{\left\|\mu_a\right\|\left\|\mu_b\right\|}\right)$	
Normalized city block	$NL_1 = \sum_{k=1}^{K} \frac{\left\|m_{ak} - m_{bk}\right\|}{\left(\sqrt{c_{ak}} + \sqrt{c_{bk}}\right)/2}$	Extension simple measures of separation of the means and covariance.
Mahalanobis	$MH = \left[\left(\mu_a - \mu_b\right)^T \left(\frac{C_a + C_b}{2}\right)^{-1} \left(\mu_a - \mu_b\right)\right]^{1/2}$	
Divergence	$D = \frac{1}{2} tr\left[\left(C_a - C_b\right)\left(C_b^{-1} - C_a^{-1}\right)\right]$ $+ \frac{1}{2} tr\left[\left(C_a^{-1} - C_b^{-1}\right)\left(\mu_a - \mu_b\right)\left(\mu_a - \mu_b\right)^T\right]$	Assume the distribution of class a and b are normal and using both means and covariances in calculation.
Transformed divergence	$D^t = 2\left[1 - e^{-D/8}\right]$	
Bhattacharyya	$B = \frac{1}{8} MH + \frac{1}{2} \ln\left[\frac{\left\|(C_a + C_b)/2\right\|}{\left(\left\|C_a\right\|\left\|C_b\right\|\right)^{1/2}}\right]$	
Jeffries-Matusita	$JM = \left[2(1 - e^{-B})\right]^{1/2}$	

Modified from Schowengerdt (1997)

According to Jensen (2005), the transformed divergence "gives an exponentially decreasing weight to increasing distances between the classes." The scale of the divergence values can range from 0 to 2,000. Interpreting your results after applying transformed divergence requires you to analyze those numerical divergence values. As a general rule, if the result is greater than 1,900, then the classes can be separated. Between 1,700 and 1,900, the separation is fairly good. Below 1,700, the separation is poor.

In addition, contingency matrix among the training dataset can be evaluated homogeneity. In fact, the pixels of each training sample are not always so homogeneous that every pixel in a sample is actually classified to its corresponding class. Each sample pixel only weights the statistics that determine the classes. However, if the signature statistics for each sample are distinct from those of the other samples, then a high percentage of each sample's pixels is classified as expected. In this evaluation, a quick classification of the sample pixels is performed based on selected decision rule such as minimum distance, maximum likelihood, or Mahalanobis distance. ERDAS (2002) mentioned that a contingency matrix contained the number and percentages of pixels that are classified as expected.

7.5 Image Classification Algorithm

In this chapter, common image classification algorithms that are applied for multispectral classification will be here separately explained in three groups: parametric, non-parametric and non-metric classification algorithms in the following sections.

7.5.1 Parametric Classification

Parametric algorithms rely on assumption about the form of the probability distribution for each class. The most notable example is maximum-likelihood, which explicitly uses a probability model to determine the decision boundaries. The necessary parameters for the model are estimates from training data. A useful property of parametric classifiers, that is not possible with nonparametric classifiers, is theoretical estimation of classifiers error from the assumed distribution (Schowengerdt, 1997). A parametric decision rule is trained by the parametric signatures. These signatures are defined by the mean vector and covariance matrix for the data file values of th pixels in the signatures. When a parametric decision rule is used, every pixel is assigned to a cla since the parametric decision space is continuous (Kloer, 1994).

The most commonly used parametric classification algorithms include:

- Maximum likelihood classifier
- Clustering classifier

7.5.1.1 Maximum Likelihood Classifiers

The maximum likelihood decision rule is based on probability. It assigns each pixel having pattern measurements or features X to the class i whose units are most probable or likely to have given rise to feature vector X. In other words, the probability of a pixel belonging to each of a predefined set of m classes is calculated, and the pixel is then assigned to the class for which the probability is the highest. The maximum likelihood decision rule is still one of the most widely used supervised classification algorithms (Jensen, 2005).

ERDAS (2002) stated that the maximum likelihood decision rule is based on the probability that a pixel belongs to a particular class. The basic equation assumes that these probabilities are equal for all classes, and that the input bands have normal distributions.

The maximum likelihood procedure assumes that the training data statistics for each class in each band are normally distributed (Gaussian). Training data with bi- or n-modal histograms in a single band are not ideal. In such cases the individual modes probably represent unique classes that should be trained upon individually and labeled as separate training classes. This should then produce unimodal, Gaussian training class statistics that fulfill the normal distribution requirement.

The computation of an n-dimensional multivariate normal density function for the classes of interested using the equation:

$$p(X \mid w_i) = \frac{1}{(2\pi)^{\frac{n}{2}} |V_i|^{\frac{1}{2}}} \exp\left[-\frac{1}{2} (X - M_i)^T V_i^{-1} (X - M_i) \right] \quad (7\text{-}1)$$

Where:

$p(X \mid w_i)$ is the probability density function of class w_i

$|V_i|$ is the determinant of the covariance matrix,

V_i^{-1} is the inverse of the covariance matrix,

$(X - M_i)^T$ is the transpose of the vector $(X - M_i)$.

The mean vectors (M_i) and covariance matrix (V_i) for each class are estimated from the training data.

If we assume that there are m classes, then p(X|wi) is the probability density function associated with the unknown measurement vector X, given that X is from a pattern in class w_i. In this case the maximum likelihood decision rule becomes:

Decide X ∈ wi, if, and only if,

$$p(X \mid w_i) \cdot p(w_i) \geq p(X \mid w_j) \cdot p(w_j) \qquad (7\text{-}2)$$

for all i and j out of 1, 2, ... m possible classes.

Therefore, to classify a pixel in the multispectral remote sensing dataset with an unknown measurement vector X, a maximum likelihood decision rule computes the product p(X|wi) · p(wi) for each class and assigns the pattern to the class having the largest product. This assumes that we have some useful information about the prior probabilities of each class i (Jensen, 2005).

A. Maximum Likelihood Classification without Prior Probability Information

In practice, we rarely have prior information about whether one class (e.g., forest) is expected to occur more frequently in a scene than any other class (e.g., 60% of the scene should be forest). This is called class a priori probability information (i.e., p(wi)). Therefore, most applications of the maximum likelihood decision rule assume that each class has an equal probability of occurring in the landscape. This makes it possible to remove the prior probability term (p(wi)) in Equation 9-13 and develop a simplified decision rule that can be applied to the unknown measurement vector X for each pixel in the scene:

Decide unknown measurement vector X is in class i if, and only if,

$p_i >= p_j$ (7-3)

for all i and j out of 1, 2, ... m possible classes

and:

$$p_i = \frac{1}{2}\log_e |V_i| - \left[\frac{1}{2}(X - M_i)^T V_i^{-1}(X - M_i)\right] \qquad (7\text{-}4)$$

Where:

Mi is the mean measurement vector for class I,

Vi is the covariance matrix of class i for bands k through l.

Therefore, to assign the measurement vector X of an unknown pixel to a class, the maximum likelihood decision rule computes the value pi for each class. Then it assigns the pixel to the class that has the largest (or maximum) value. This assumes that we have no useful information about the prior probabilities of each class, that is, every class has an equal probability of occurring in the landscape (Richards and Jia, 1999).

An example of maximum likelihood classification without prior probability information was presented by Jensen (2005) as following:

In the first pass, p_1 is computed with V_1 and M_1 being the covariance matrix and mean vectors for class 1. Next, p_2 is computed with V_2 and M_2 being the covariance matrix and mean vectors for class 2. This continues for all m classes. The pixel or measurement vector X is assigned to the class that produces the largest or maximum pi. As shown in Figure 7-8, pixel X would be assigned to forest because the probability density of unknown measurement vector X is greater for greater for forest than agriculture.

B. Maximum Likelihood Classification with Prior Probability Information

Equation 7-3 assumes that each class has an equal probability of occurring in the terrain. However, common sense reminds us that in most remote sensing applications there is a high probability of encountering some classes more often than others. For examples the probabilities of major land use classes are:

p(w1) = p(residential) = 0.2,
p(w2) = p(commercial) = 0.1,
p(w3) = p(wetland) = 0.25,
p(w4) = p(forest) = 0.1,
p(w5) = p(water) = 0.35.

Thus, we would expect more pixels to be classified as water because it is more prevalent in the terrain. If we have such information, it is possible to include this variable prior knowledge in the classification decision. We can do this by weighting each class i by its appropriate a priori probability (p(wi) (Jensen, 2005). The equation then becomes:

Decide unknown measurement vector X is in class i if, and only if,

$$p_i \cdot p(w_i) \geq p_j \cdot p(w_j) \qquad (7\text{-}5)$$

for all i and j out of 1,2,... m possible classes, and

$$p_i \cdot p(w_i) = \log_e p(w_i) - \frac{1}{2}\log_e |V_i| - \left[\frac{1}{2}(X - M_i)^T V_i^{-1}(X - M_i)\right] \qquad (7\text{-}6)$$

This Bayes's decision rule is identical to the maximum likelihood decision rule (Eq. 7-4) except that it does not assume that each class has equal probability. The maximum likelihood and Bayes's classification require many more computations per pixel than either the parallelepiped or minimum distance classification algorithms. They do not always produce superior results. Advantage and disadvantage of maximum likelihood classifier was shown in Table 7-3.

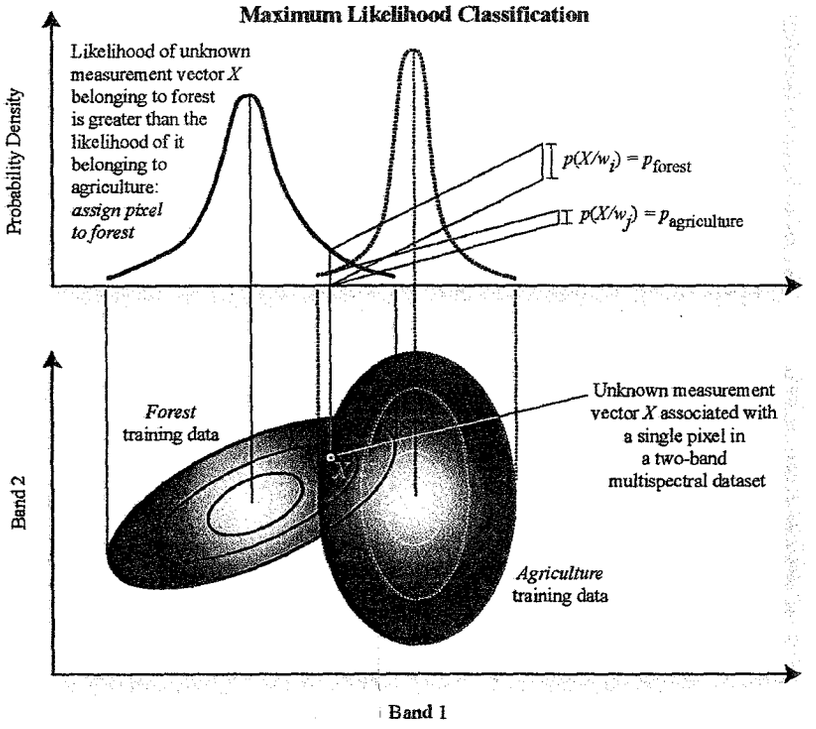


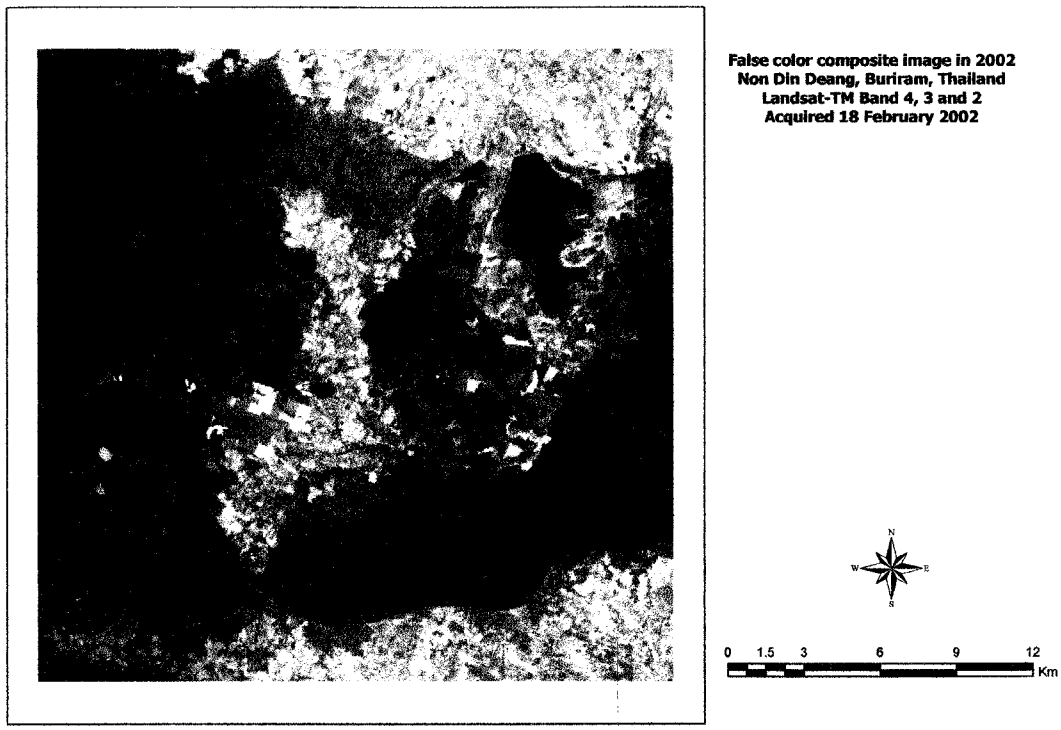
Figure 7-8: How the maximum likelihood decision rule function (From, Jensen, 2005)

Table 7-3: Advantage and disadvantage of maximum likelihood decision rule.

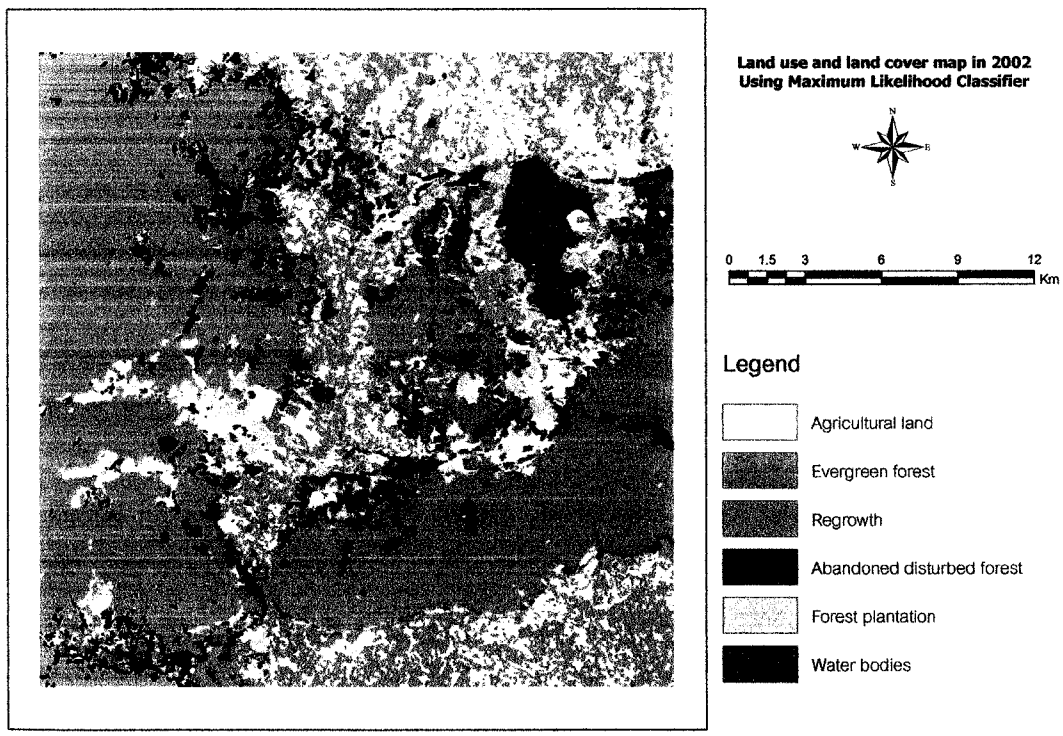
Advantages	**Disadvantages**
The most accurate of the classifiers in the ERDAS IMAGINE system (if the input samples/clusters have a normal distribution), because it takes the most variables into consideration.	An extensive equation that takes a long time to compute. The computation time increases with the number of input bands.
Takes the variability of classes into account by using the covariance matrix, as does Mahalanobis distance.	Maximum likelihood is parametric, meaning that it relies heavily on a normal distribution of the data in each input band.
	Tends to overclassify signatures with relatively large values in the covariance matrix.

Form ERDAS, 2002.

Example of land use and land cover classification using maximum likelihood classifiers was shown in Figure 7-9.



(a) False color composite data of Landsat-TM: Don Din Deang, Buriram, Thailand, Acquiring date: 18 February 2002



(b) Land use and land cover in 2002 map : Don Din Deang, Buriram, Thailand, Using maximum likelihood decision rule

Figure 7-9: Land use and land cover classification using maximum likelihood decision rule

7.5.1.2 Clustering

Unsupervised classification that commonly referred to as clustering is an effective method of partitioning remote sensor image data in multispectral feature space and extracting land-cover information. Compared to supervised classification, unsupervised classification normally requires only a minimal amount of initial input from the analyst. This is because clustering does not normally require training data.

Unsupervised classification is the process where numerical operations are performed that search for natural groupings of the spectral properties of pixels, as examined in multispectral feature space. The clustering process results in a classification map consisting of m spectral classes. The analyst then attempts a posteriori (after the fact) to assign or transform the spectral classes into thematic information classes of interest (e.g., forest, agriculture). This may be difficult. Some spectral clusters may be meaningless because they represent mixed classes of Earth surface materials. The analyst must understand the spectral characteristics of the terrain well enough to be able to label certain clusters as specific information classes (Jensen, 2005).

Hundreds of clustering algorithms have been developed. Some common clustering algorithms which commonly used in unsupervised classification here are explained include:

- K-means clustering algorithm
- ISODATA
- RGB Clustering

A. K-means Clustering algorithm

Schowengerdt (1997) claimed that K-means algorithm is one of the common clustering methods. In the first step of the algorithm, an initial mean vector (seed or attractor) is arbitrarily specified for each of K clusters. Each pixel of the training set is then assigned to the class whose mean vector is closest to the pixel vector, thus forming the first set of decision boundaries. A new set of cluster mean vectors is then calculated from this classification, the pixels are reassigned accordingly. In each iteration, the K-means will tend to settle towards concentrations of data within their currently-assigned region of feature space. The iterations are continued until there is no significant change in pixel assignments from one iteration to the next (Figure 7-10). The criterion for ending the iterative process can be defined in terms of the net mean migration from one iteration to the next. The magnitude change of the mean vectors from iteration i-1 to iteration i was summed over all K clusters as:

$$\Delta\mu(i) = \sum_{k=1}^{K} \left| \mu_k(i) - \mu_{k-1}(i-1) \right|$$

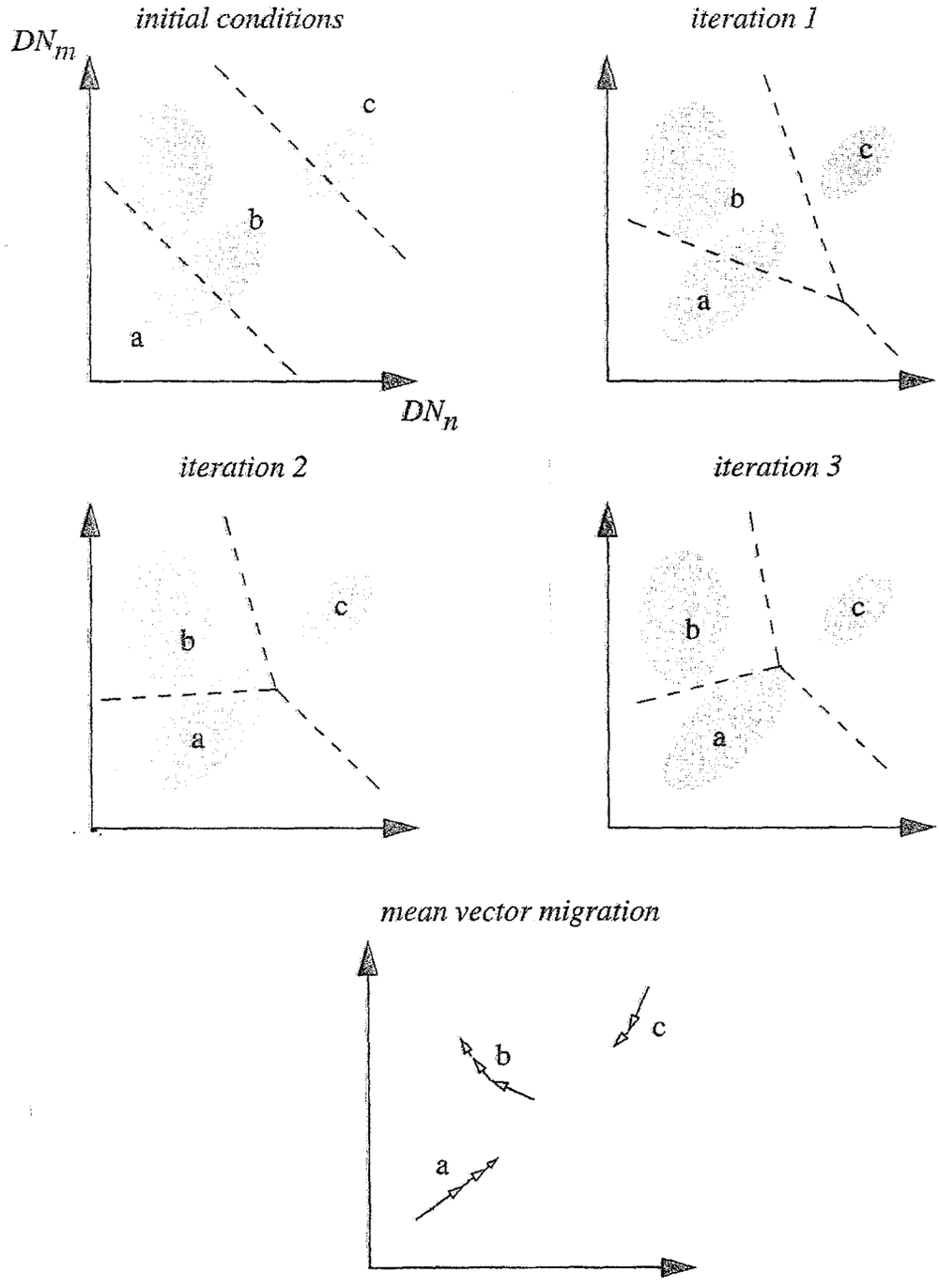


Figure 7-10: An idealized data distribution during three iterations of the K-means clustering algorithms with the nearest-mean decision criterion (From Schowengerdt, 1997)

A typical iteration sequence for the net mean migration of the cluster means was shown in Figure 7-11.

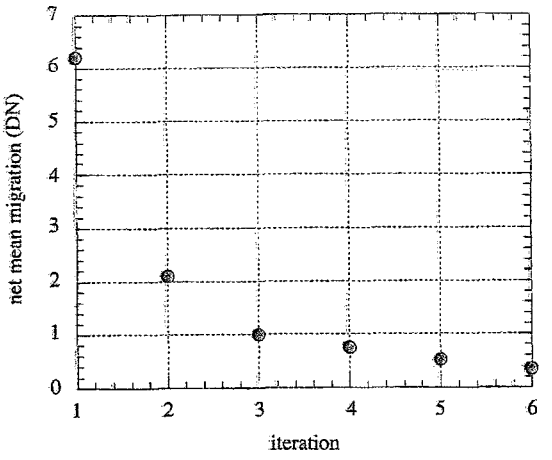


Figure 7-11: Typical behavior of the net mean migration from one iteration to the next in the K-mean algorithm. (From Schowengerdt, 1997)

B. ISODATA Algorithm

Another widely used clustering algorithm is the Iterative Self-Organizing Data Analysis Technique (ISODATA). ISODATA represents a comprehensive set of heuristic (rule of thumb) procedures that have been incorporated into an iterative classification algorithm. Many of the steps incorporated into the algorithm are a result of experience gained through experimentation (Jensen, 2005). Schowengerdt (1997) mentioned that the ISODATA algorithm is a modification of the k-means clustering algorithm, which includes:

- merging clusters if their separation distance in multispectral feature space is below a user-specified threshold and
- splitting a single cluster into two clusters.

A sophisticated ISODATA algorithm normally requires the analyst to specify the following criteria Jensen (2005):

- **C_{max}:** the maximum number of clusters to be identified by the algorithm.
- **T:** the maximum percentage of pixels whose class values are allowed to be unchanged between iterations. When this number is reached, the ISODATA algorithm terminates.
- **M:** the maximum number of times ISODATA is to classify pixels and recalculate cluster mean vectors. The ISODATA algorithm terminates when this number is reached.
- **Minimum members in a cluster (%):** If a cluster contains less than the minimum percentage of members, it is deleted and the members are assigned to an alternative cluster. The default minimum percentage of members is often set to 0.01.

- **Maximum standard deviation (σ_{max}):** When the standard deviation for a cluster exceeds the specified maximum standard deviation and the number of members in the class is greater than twice the specified minimum members in a class, the cluster is split into two clusters. The mean vectors for the two new clusters are the old class centers $\pm 1\sigma$. Maximum standard deviation values between 4.5 and 7 are typical.
- **Split separation value:** If this value is changed from 0.0, it takes the place of the standard deviation in determining the locations of the new mean vectors plus and minus the split separation value.
- **Minimum distance between cluster means (C):** Clusters with a weighted distance less than this value are merged. A default of 3.0 is often used.

The major steps of ISODATA based Jensen (2005) can be summarized as following:

(1) ISODATA Initial Arbitrary Clustering Allocation: In the first step of the ISODATA algorithm, an initial arbitrary assignment of all C_{max} clusters takes place along an n-dimensional vector that runs between very specific points in feature space. The region in feature space is defined using the mean, μk, and standard deviation, σk, of each band in the analysis (Figure 7-12a). This method of automatically seeding the original C_{max} vectors makes sure that the first few lines of data do not bias the creation of clusters.

(2) ISODATA First Iteration: With the initial C_{max} mean vectors in place, a pass is made through the database beginning in the upper left corner of the matrix. Each candidate pixel is compared to each cluster mean and assigned to the cluster whose mean is closest in Euclidean distance (Figure 7-12b). This pass creates an actual classification map consisting of C_{max} classes.

(3) ISODATA Second to M^{th} Iteration: After the first iteration, a new mean for each cluster is calculated based on the actual spectral locations of the pixels assigned to each cluster. This involves analysis of several parameters include minimum members in a cluster (%), maximum standard deviation, split separation, and minimum distance between cluster means. Then the entire process is repeated with each candidate pixel once again compared to the new cluster means and assigned to the nearest cluster mean (Figure 7-12c). This iterative process continues until there is (1) little change in class assignment between iterations (i.e., the T threshold is reached) or (2) the maximum number of iterations is reached (M). The final file is a matrix with C_{max} clusters in it (Figure 7-12d), which must be identified and labeled as land use classes. Example of land use and land cover classification using ISODATA clustering was shown in Figure 7-13.

Advantage and disadvantage of ISODATA clustering was shown in Table 7-4.

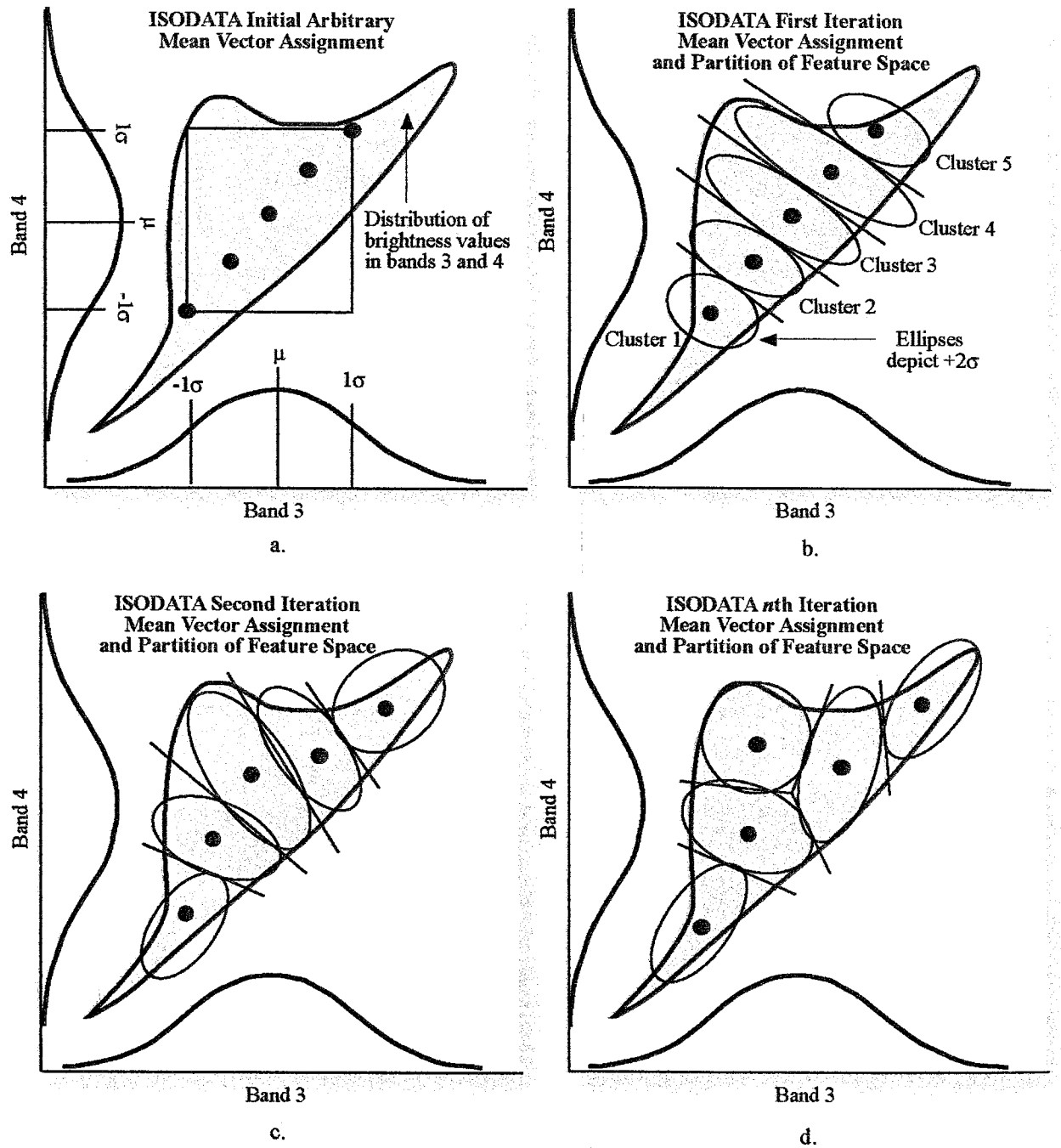
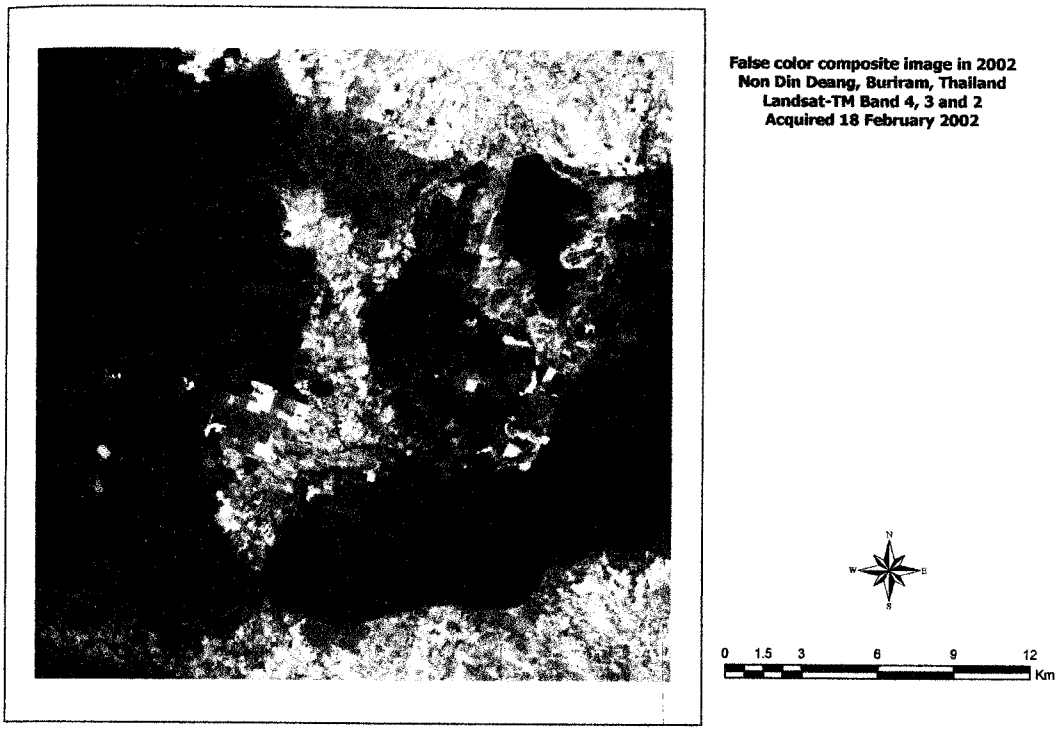
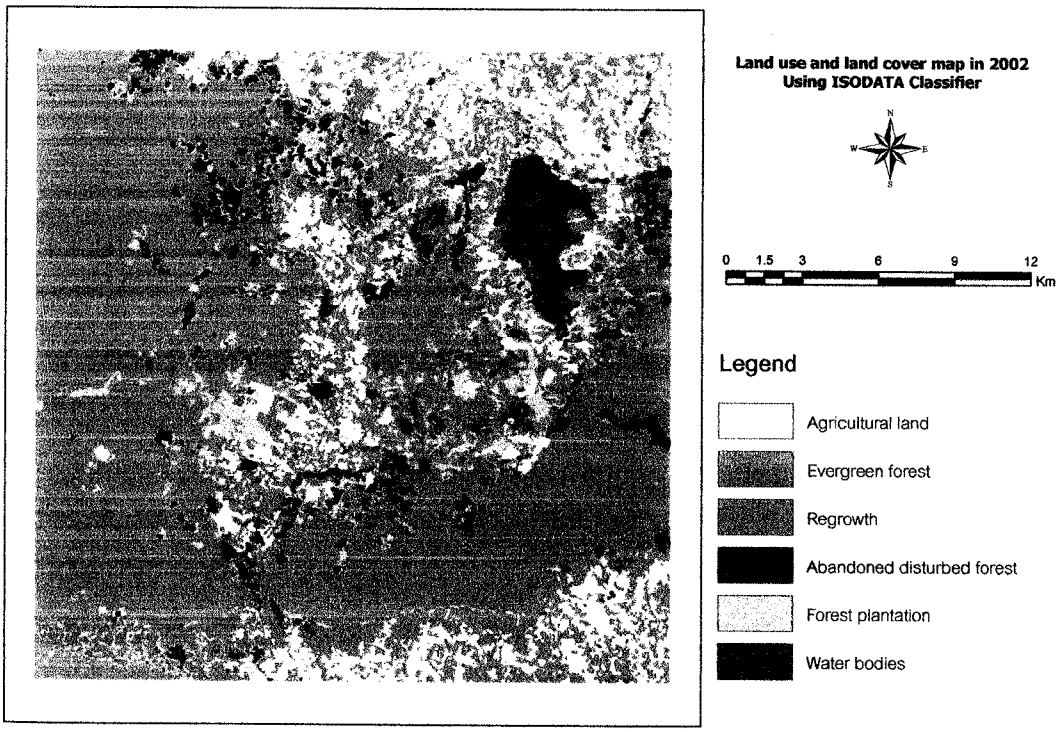


Figure 7-12: An idealized data distribution during initial stage up to n iterations of the ISODATA clustering algorithms (From Jensen, 2005)



(a) False color composite data of Landsat-TM: Don Din Deang, Buriram, Thailand, Acquiring date: 18 February 2002



(b) Land use and land cover in 2002 map : Don Din Deang, Buriram, Thailand, Using ISODATA clustering decision rule.

Figure 7-13: Land use and land cover classification using ISODATA clustering algorithms

Table 7-4: Advantage and disadvantage of ISODATA decision rule.

Advantages	Disadvantages
Because it is iterative, clustering is not geographically biased to the top or bottom pixels of the data file.	The clustering process is time-consuming, because it can repeat many times.
This algorithm is highly successful at finding the spectral clusters that are inherent in the data. It does not matter where the initial arbitrary cluster means are located, as long as enough iterations are allowed.	Does not account for pixel spatial homogeneity.
A preliminary thematic raster layer is created, which gives results similar to using a minimum distance classifier (as explained below) on the signatures that are created. This thematic raster layer can be used for analyzing and manipulating the signatures before actual classification takes place.	

From: ERDAS, 2002.

C. RGB Clustering

RGB clustering is a simple classification and data compression technique for three bands of data. It is a fast and simple algorithm that quickly compresses a three-band image into a single band pseudocolor image, without necessarily classifying any particular features. In practice, RGB Clustering differs greatly from the other clustering methods, but it does employ a clustering algorithm (ERDAS, 2002).

The algorithm plots all pixels in 3-dimensional feature space and then partitions this space into clusters on a grid. ERDAS (2002) divided RGB clustering in two versions: simple and advance RGB clustering. In the simple version, each of these clusters becomes a class in the output thematic raster layer. The advanced version requires that a minimum threshold on the clusters be set so that only clusters at least as large as the threshold become output classes. This allows for more color variation in the output file. Pixels that do not fall into any of the remaining clusters are assigned to the cluster with the smallest city-block distance from the pixel. In this case, the city-block distance is calculated as the sum of the distances in the red, green, and blue directions in 3-dimensional space.

is calculated as the sum of the distances in the red, green, and blue directions in 3-dimensional space.

Theoretically, along each axis of the three-dimensional scatterplot, each input histogram is scaled so that the partitions divide the histograms between specified limits—either a specified number of standard deviations above and below the mean, or between the minimum and maximum data values for each band as shown in Figure 7-14. The default number of divisions per band is listed below:

- Red is divided into 7 sections (32 for advanced version)
- Green is divided into 6 sections (32 for advanced version)
- Blue is divided into 6 sections (32 for advanced version)

It is necessary to specify the number of R, G, and B sections in each dimension of the 3-dimensional scatterplot. The number of sections should vary according to the histograms of each band. Broad histograms should be divided into more sections, and narrow histograms should be divided into fewer sections. Starting values that usually produce good results with the simple RGB clustering are: R = 7, G = 6 and B = 6, which results in $7 \times 6 \times 6 = 252$ classes. To decrease the number of output colors/classes or to darken the output, decrease these values. For the Advanced RGB clustering function, start with higher values for R, G, and B. Adjust by raising the threshold parameter and/or decreasing the R, G, and B parameter values until the desired number of output classes is obtained (ERDAS, 2002). The advantage and disadvantage of RGB clustering was summarized in Table 7-5.

Example of land use and land cover classification using RGB clustering algorithms was shown in Figure 7-15.

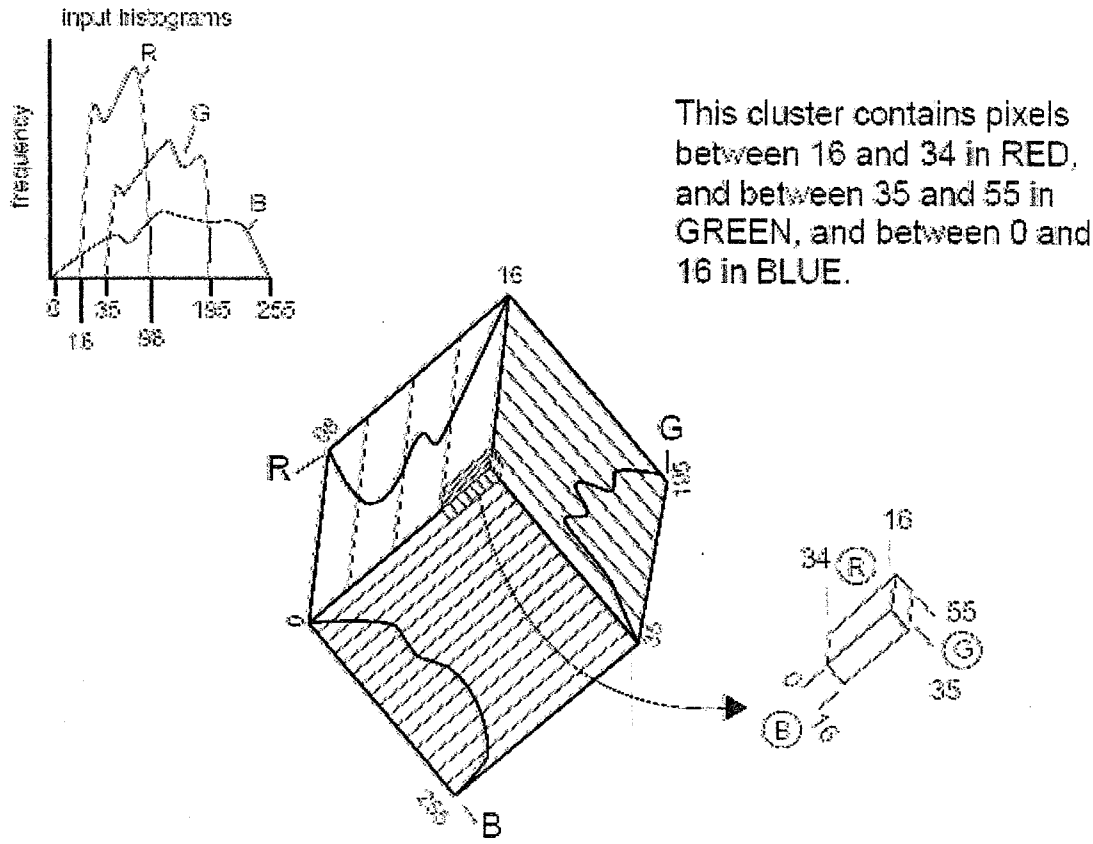
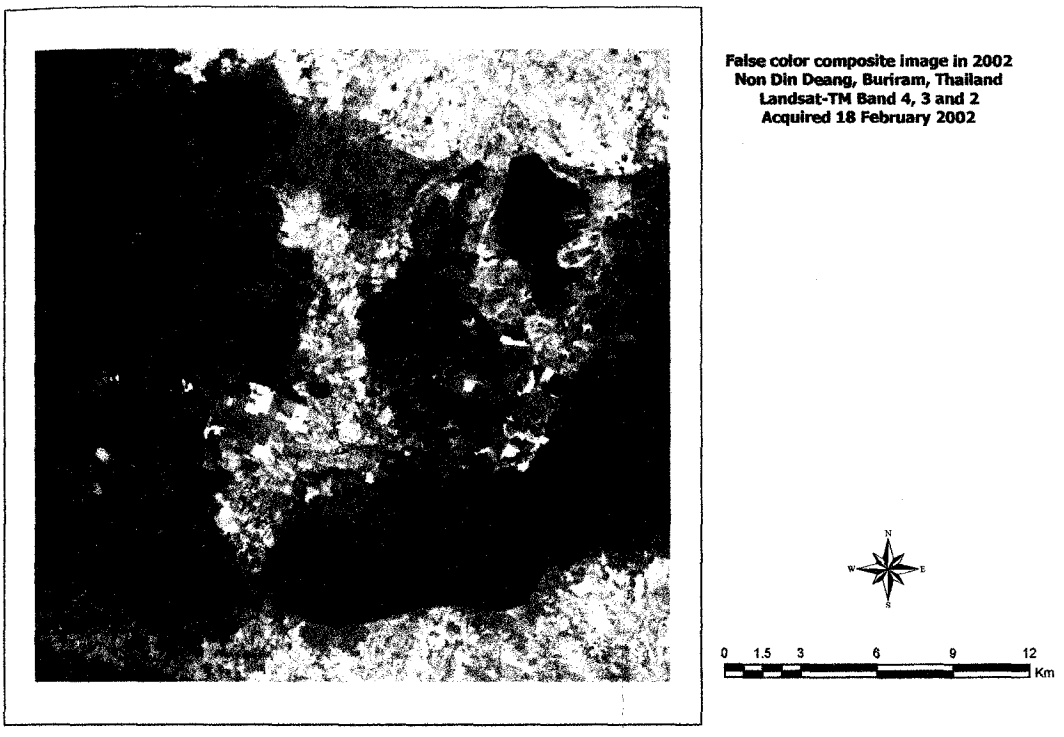


Figure 7-14: RGB clustering algorithm (From ERDAS, 2002)

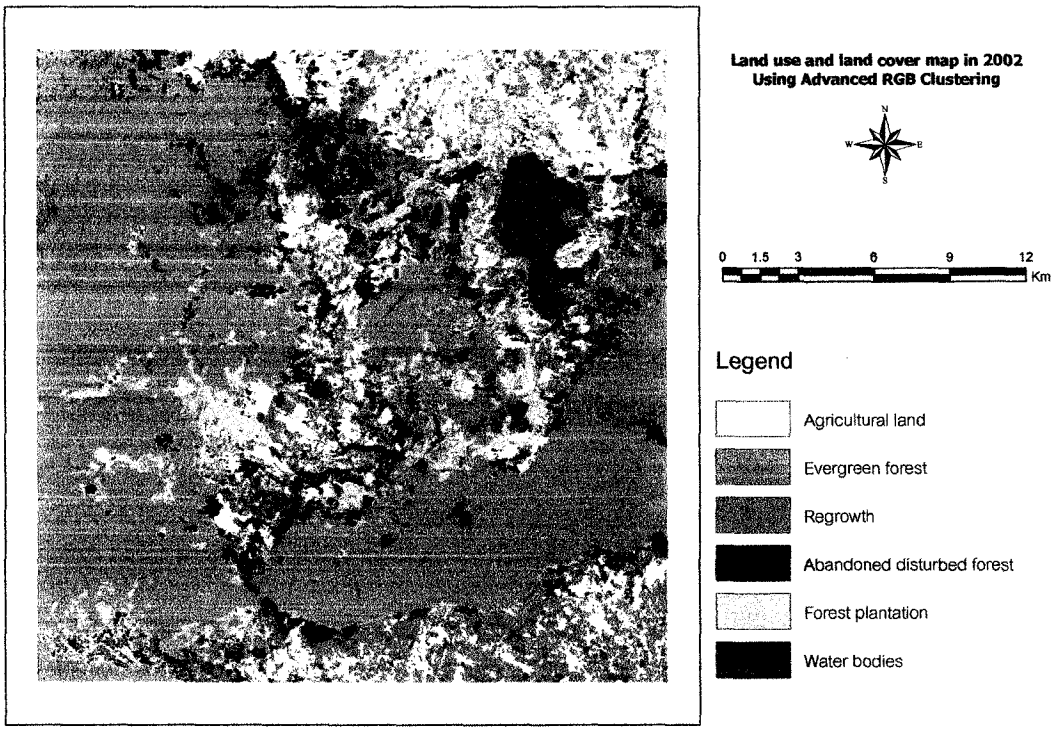
Table 7-5: Advantage and disadvantage of RGB clustering

Advantages	**Disadvantages**
The fastest classification method. It is designed to provide a fast, simple classification for applications that do not require specific classes.	Exactly three bands must be input, which is not suitable for all applications.
Not biased to the top or bottom of the data file. The order in which the pixels are examined does not influence the outcome.	Does not always create thematic classes that can be analyzed for informational purposes.
A highly interactive function, allowing an iterative adjustment of the parameters until the number of clusters and the thresholds are satisfactory for analysis.	

From ERDAS, 2002.



(a) False color composite data of Landsat-TM: Don Din Deang, Buriram, Thailand, Acquiring date: 18 February 2002



(b) Land use and land cover in 2002 map : Don Din Deang, Buriram, Thailand, Using Advanced RGB clustering decision rule.

Figure 7-15: Land use and land cover classification using RGB clustering algorithms

7.5.2 Nonparametric Classification

Nonparametric algorithms make no assumption about the probability distribution and are often considered robust because they may work well for a variety of classification, as long as the class signatures are reasonably distinct (Showengerdt, 1997). ERDAS (2002) state that a nonparametric decision rule is not based on statistics; therefore, it is independent of the properties of the data. If a pixel is located within the boundary of a nonparametric signature, then this decision rule assigns the pixel to the signature's class. Basically, a nonparametric decision rule determines whether or not the pixel is located inside of nonparametric signature boundary.

The most commonly used nonparametric classification algorithms included:

- Level-slice Classifier
- Parallelepiped Classifier
- Minimum distance Classifier
- Nearest neighbor Classifier
- Fuzzy Classifier
- Artificial Neural Network

7.5.2.1 Level-Slice Classifier

This classifier, also known as a box classifier, is perhaps the simplest of all classification methods. A set of k-dimensional boxes, centered at the estimated class mean vectors, are placed in k-dimensional feature space. If an unlabeled pixel vector lies within one of the boxes, it is assigned that class label (Figure 7-16). The level-slice algorithm also creates an unlabeled class. This characteristic can be used to prevent overestimation of class populations caused by outliers (Schowengerdt, 1997).

Specification of the box limits is typically in terms of the data extent in each dimension, for example ± one standard deviation about the mean. However, a complication occurs if a pixel vector falls within two or more boxes. A decision on a pixel's label must then be made with another algorithm, such as the nearest-mean.

Given the class DN distributions are not typically aligned with the k-dimensional data axes because of correlation between spectral bands. A modification, the parallelepiped classifier, allows hyperdimensional boxes that align with the data clusters and whose faces are parallelograms rather than rectangles (Schowengerdt, 1997).

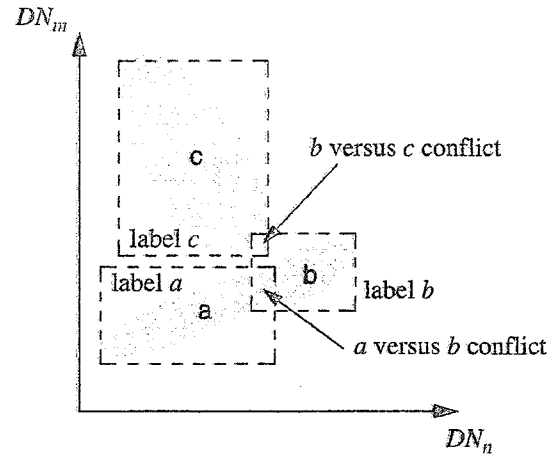


Figure 7-16: Level-slice decision boundaries for three classes in two dimensions. (From Schowengerdt, 1997)

7.5.2.2 Parallelepiped Classifier

This is a modified of level-slicing classifiers and a widely used digital image classification decision rule based on simple Boolean "and/or" logic. Training data in n spectral bands are used to perform the classification. Brightness values from each pixel of the multispectral imagery are used to produce an n-dimensional mean vector, Mc = (μ_{ck}, μ_{c2}, μ_{c3}, ..., μ_{cn}) with μ_{ck} being the mean value of the training data obtained for class c in band k out of m possible classes and σ_{ck} is the standard deviation of the training data class c of band k out of m possible classes.

Using a one-standard deviation threshold, a parallelepiped algorithm decides BV_{ijk} is in class c if, and only if:

$$\mu_{ck} - \sigma_{ck} \leq BV_{ijk} \leq \mu_{ck} + \sigma_{ck} \quad (7\text{-}7)$$

where

c = 1, 2, 3, ..., m, number of classes, and

k = 1, 2, 3, ..., n, number of bands.

Therefore, if the low and high decision boundaries are defined as:

$$L_{ck} = \mu_{ck} - \sigma_{ck} \quad (7\text{-}8)$$

and

$$H_{ck} = \mu_{ck} + \sigma_{ck} \quad (7\text{-}9)$$

the parallelepiped algorithm becomes

$$L_{ck} \leq BV_{ijk} \leq H_{ck} \quad (7\text{-}10)$$

These decision boundaries form an n-dimensional parallelepiped in feature space. If the pixel value lies above the low threshold and below the high threshold for all n bands evaluated, it is assigned to that class (Jensen, 2005). Figure 7-17 shows parallelepiped classification using ± two Standard deviations as limits.

The parallelepiped algorithm is a computationally efficient method of classifying remote sensor data. Unfortunately, because some parallelepipeds overlap, it is possible that an unknown candidate pixel might satisfy the criteria of more than one class. In such cases it is usually assigned to the first class for which it meets all criteria. A more elegant solution is to take this pixel that can be assigned to more than one class and use a minimum distance to mean decision rule to assign it to just one class. Example of land use and land cover classification using parallelepiped algorithm was shown in Figure 7-18. ERDAS (2002) summarized advantage and disadvantage of parallelepiped classifier as shown in Table 7-6.

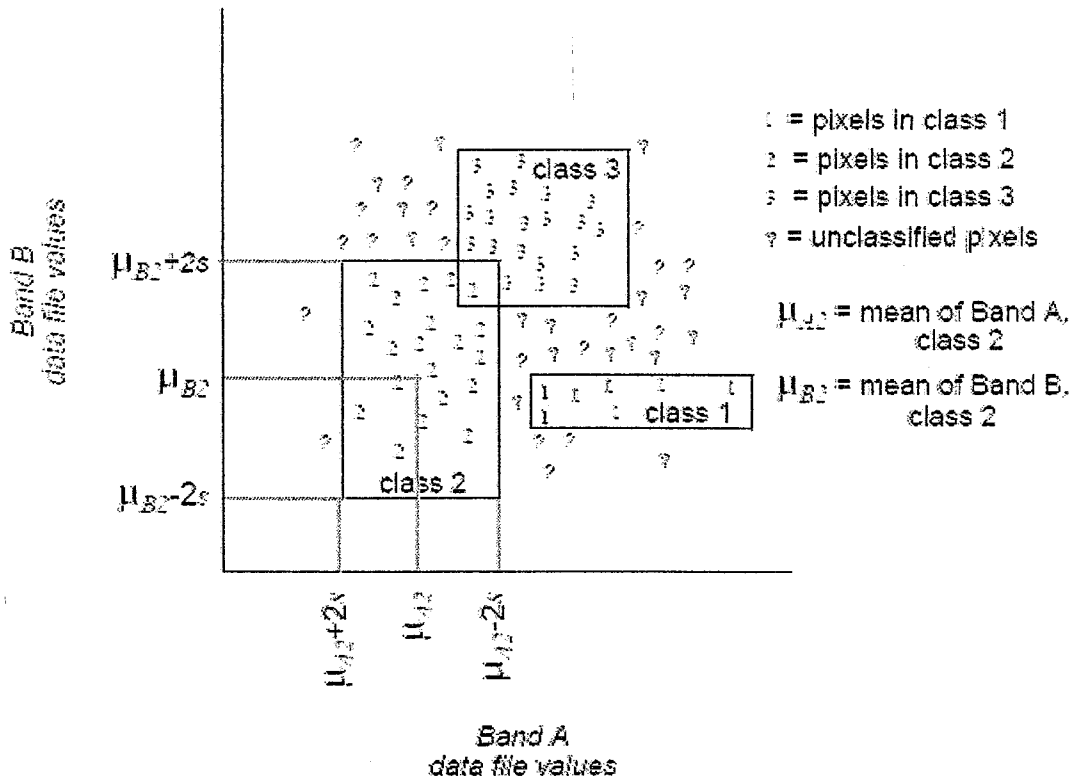
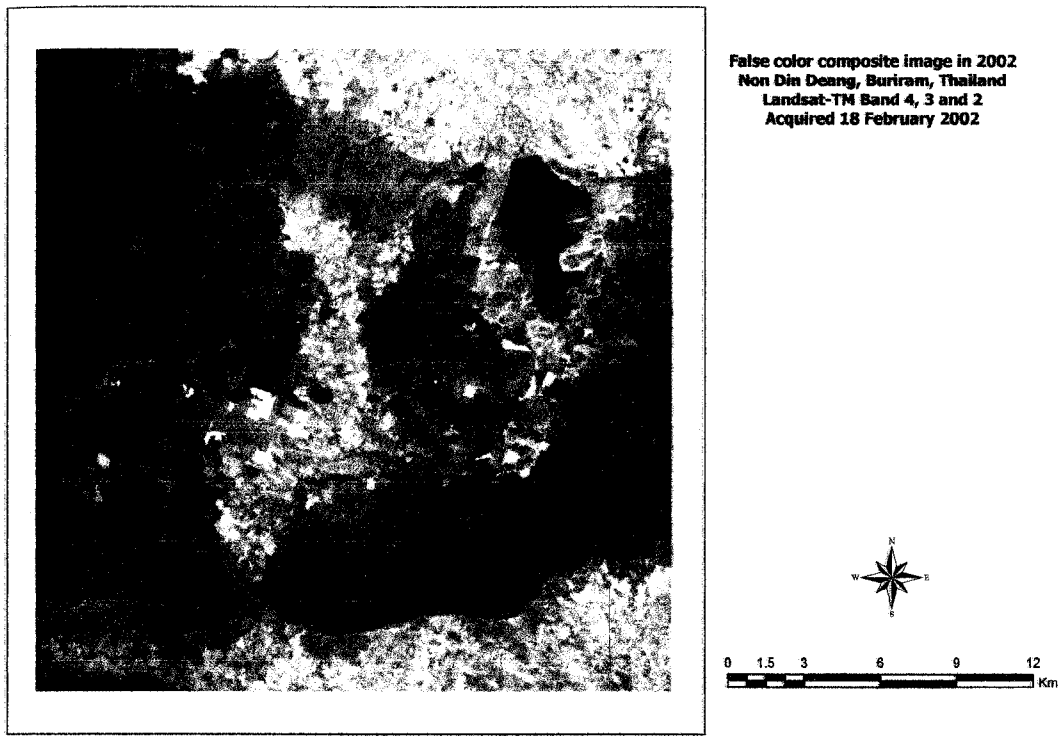
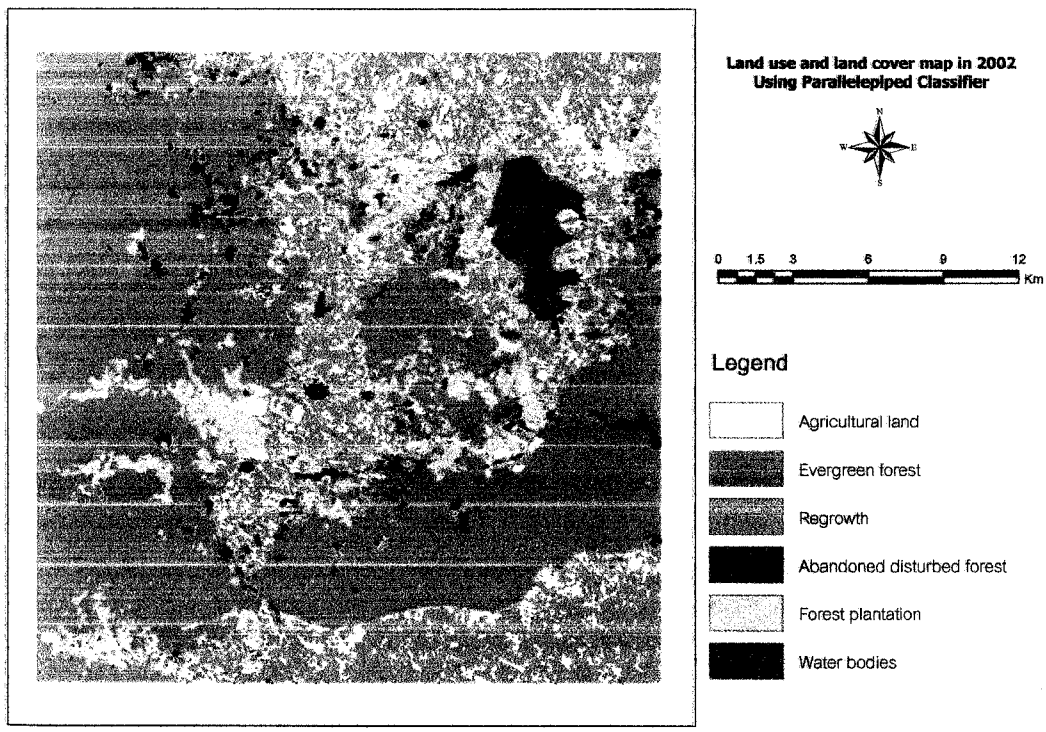


Figure 7-17: Parallelepiped classification using ± two Standard deviations as limits. (From ERDAS, 2002)



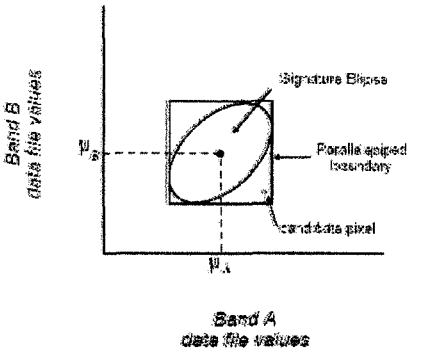
(a) False color composite data of Landsat-TM: Don Din Deang, Buriram, Thailand, Acquiring date: 18 February 2002



(b) Land use and land cover in 2002 map : Don Din Deang, Buriram, Thailand, Using parallelepiped decision rule

Figure 7-18: Land use and land cover classification using parallelepiped decision rule

Table 7-6: Advantage and disadvantage of parallelepiped decision rule.

Advantages	Disadvantages
Fast and simple, since the data file values are compared to limits that remain constant for each band in each signature.	Since parallelepipeds have corners, pixels that are actually quite far, spectrally, from the mean of the signature may be classified as shown in Figure below. 
Often useful for a first-pass, broad classification, this decision rule quickly narrows down the number of possible classes to which each pixel can be assigned before the more time-consuming calculations are made, thus cutting processing time (e.g., minimum distance, Mahalanobis distance, or maximum likelihood).	
Not dependent on normal distributions.	

From ERDAS, 2002

7.5.2.3 Minimum Distance to Means Classifier

The minimum distance to means decision rule is computationally simple and commonly used. When used properly it can result in classification accuracy comparable to other more computationally intensive algorithms such as the maximum likelihood algorithm. Like the parallelepiped algorithm, it requires that the user provide the mean vectors for each class in each band μck from the training data. To perform a minimum distance classification, a program must calculate the distance to each mean vector (μ_{ck}) from each unknown pixel (BV_{ijk}) as shown Figure 7-19. It is possible to calculate this distance using Euclidean distance based on the Pythagorean theorem or "round the block" distance measures (Jensen, 2005) as shown in Figure 7-20.

Many minimum distance algorithms let the analyst specify a distance or threshold from the class means beyond which a pixel will not be assigned to a category even though it is nearest to the mean of that category.

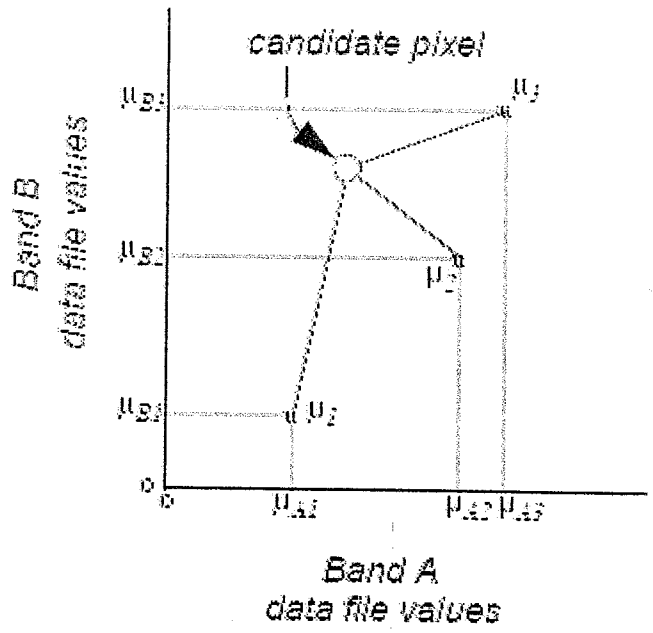


Figure 7-19: Calculation of spectral distance to means (From ERDAS, 2002)

Swain and Davis (1978) suggested the equation for classifying by spectral distance is based on the equation for Euclidean distance as:

$$SD_{xyc} = \sqrt{\sum_{i=1}^{n} \left(\mu_{ci} - X_{xyi} \right)^2} \qquad (7\text{-}11)$$

Where:

n is number of bands (dimensions)

i is a particular band

c is a particular class

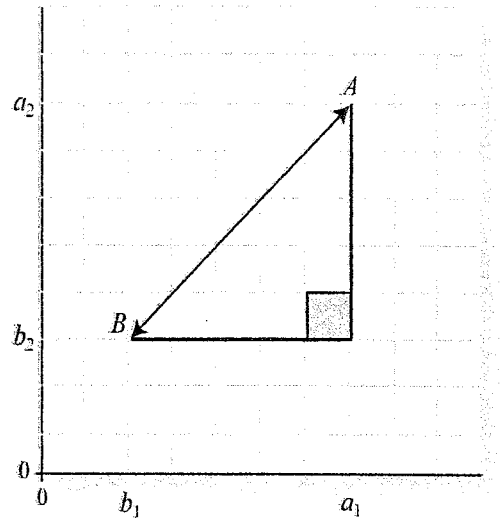
X_{xyi} is data file value of pixel x,y in band i

M_{ci} is mean of data file values in band i for the sample for class c

SD_{xyc} is spectral distance from pixel x,y to the mean of class c

Thus, when spectral distance is computed for all possible values of c (all possible classes), the class of the candidate pixel is assigned to the class for which SD is the lowest.

Advantage and disadvantage of minimum distance decision rule was summarized in Table 7-7 (ERDAS, 2002). Example of land use and land cover classification using minimum distance decision rule was shown in Figure 7-18.



Euclidean distance

$$D_{AB} = \sqrt{\sum_{i=1}^{2} (a_i - b_i)^2}$$

Round the block distance

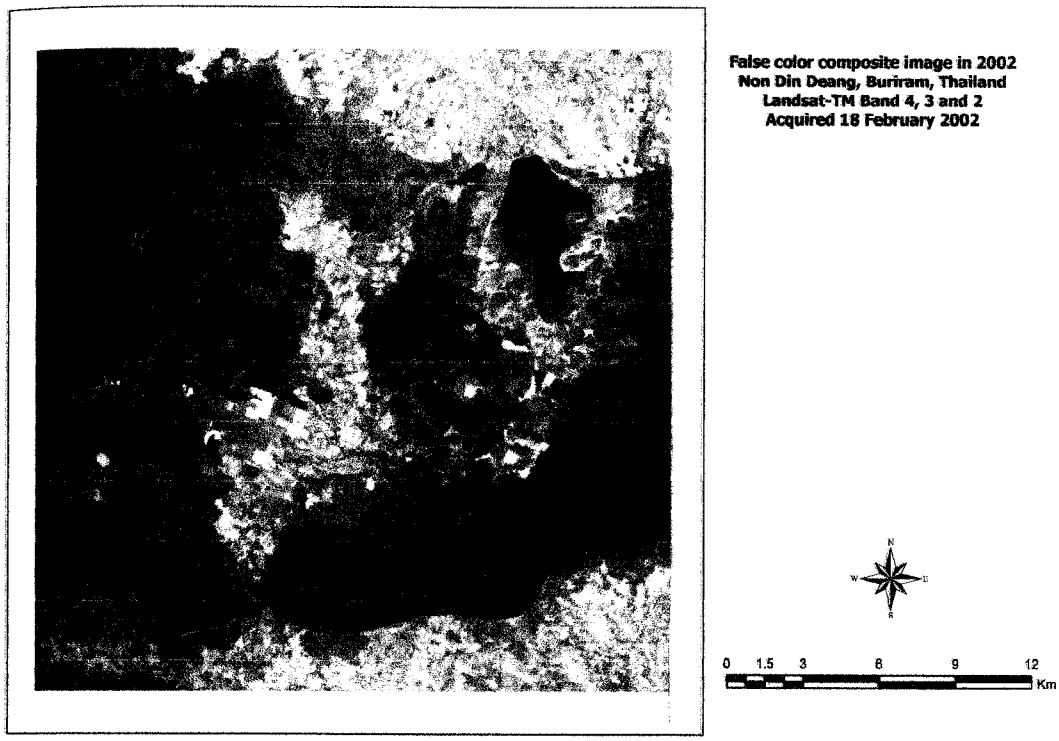
$$D_{AB} = \sum_{i=1}^{2} |(a_i - b_i)|$$

Figure 7-20: Distance use in a minimum distance to means classification algorithm (Jensen,2005)

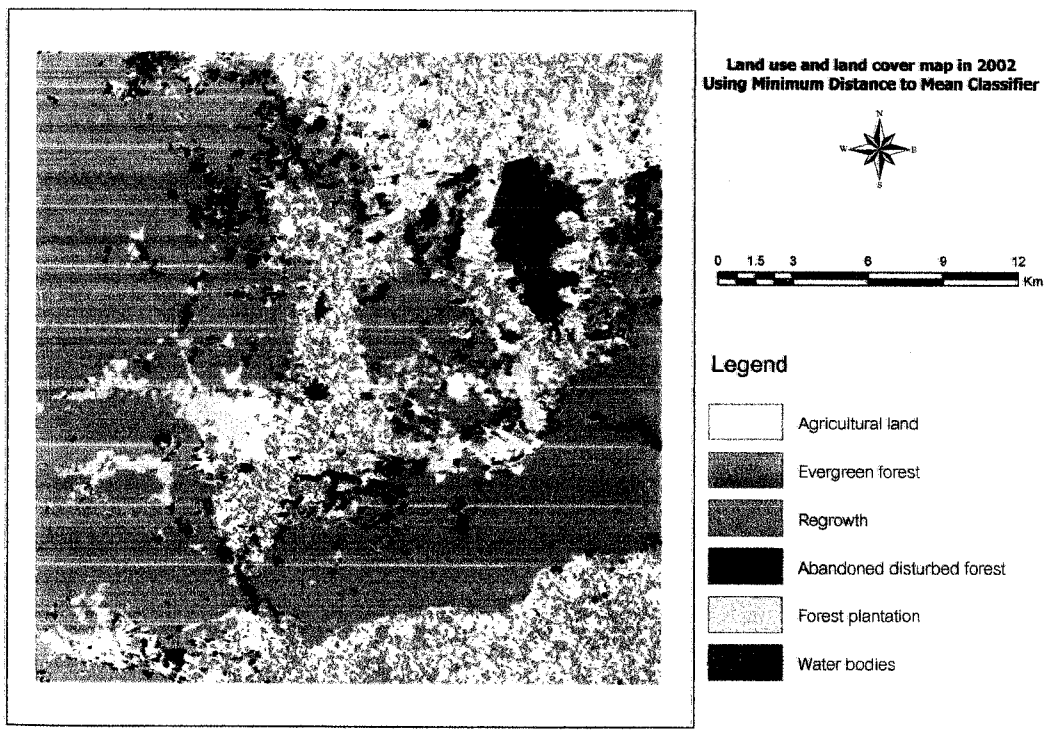
Table 7-7: Advantage and disadvantage of minimum distance decision rule

Advantages	**Disadvantages**
Since every pixel is spectrally closer to either one sample mean or another, there are no unclassified pixels.	Pixels that should be unclassified (i.e., they are not spectrally close to the mean of any sample, within limits that are reasonable to you) become classified. However, this problem is alleviated by thresholding out the pixels that are farthest from the means of their classes.
The fastest decision rule to compute, except for parallelepiped.	Does not consider class variability. For example, a class like an urban land cover class is made up of pixels with a high variance, which may tend to be farther from the mean of the signature. Using this decision rule, outlying urban pixels may be improperly classified. Inversely, a class with less variance, like water, may tend to overclassify (that is, classify more pixels than are appropriate to the class), because the pixels that belong to the class are usually spectrally closer to their mean than those of other classes to their means.

From ERDAS, 2002.



(a) False color composite data of Landsat-TM: Don Din Deang, Buriram, Thailand, Acquiring date: 18 February 2002



(b) Land use and land cover in 2002 map : Don Din Deang, Buriram, Thailand, Using Minimum distance decision rule

Figure 7-21: Land use and land cover classification by using Minimum distance decision rule

7.5.2.4 Nearest-Neighbor Classifiers

The parallelepiped classifier made use of the training class mean and standard deviation statistics for each band. The minimum distance classifiers required the mean for each training class for each band. It is also possible to classify an unknown pixel measurement vector into m classes using just the training data brightness value in each band and nearest-neighbor distance measurement (Jensen, 2005). The most common nonparametric nearest-neighbor classifiers are:

- Nearest neighbor,
- k-nearest neighbor, and
- k-nearest neighbor distance-weighted.

A. Nearest neighbor: The simple nearest-neighbor classifier computes the Euclidean distance from the pixel to be classified to the nearest training data pixel in n-dimensional feature space and assigns it to the class.

B. K-nearest-neighbor classifier: The k-nearest-neighbor classifier searches away from the pixel to be classified in all directions until encounter k user-specified training pixel (e.g., k = 5). It then assigns the pixel to the class with the majority of pixels encountered.

C. K-nearest-neighbor distance-weighted classifier: The k-nearest-neighbor distance-weighted classifier uses the same k-nearest neighbors but weight them according to the distance-weight logic. The pixel under investigation is assigned to the training class pixel with the highest total weight.

Hypothetical example of nearest-neighbor classification was shown Figure 7-22. Nearest-neighbor classifiers can be relatively slow because of the number of distance calculations required between the unknown pixel measurement vector and all of the training pixels in the various bands. Nearest-neighbor classifiers can yield useful results if the training data are well separated in n - dimensional feature space. Otherwise, it will probably by necessary to use a different algorithm (Jensen, 2005).

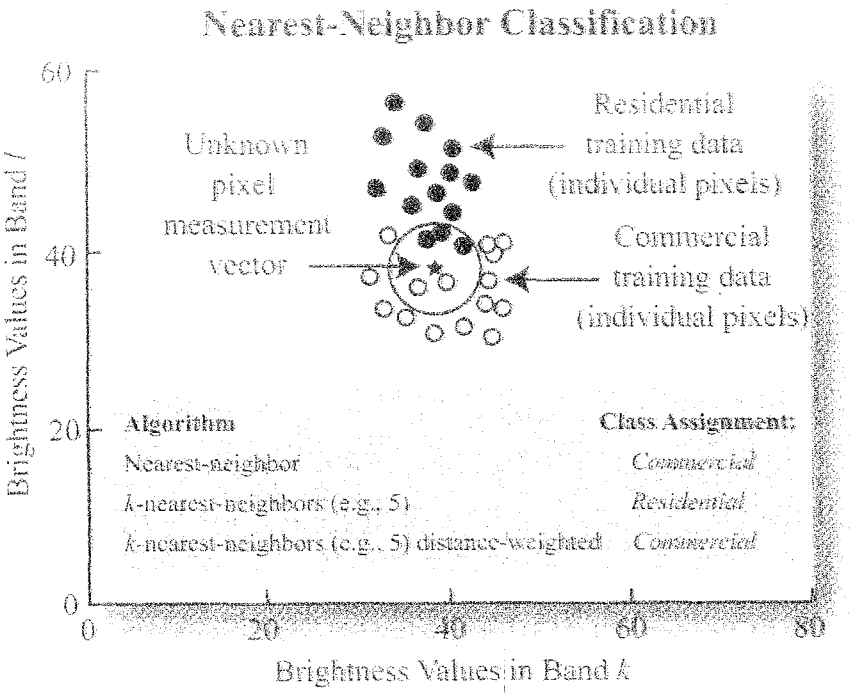


Figure 7-22: Hypothetical example of nearest-neighbor classification (From Jensen, 2005).

7.5.2.5 Sub-Pixel Classification

A remotely sensed image is simply a representation of the way radiation interacts with the Earth's surface. As a representation of the environment some degree of inconsistency with reality is to be expected. Various sources of inconsistency may be identified but of particular significance to the topic of this chapter is a basic property of the remotely sensed imagery. Typically the imagery comprises raster data and the fundamental spatial unit, the pixel, is defined primarily by sensor dependent variables. That is, the pixel is an arbitrary spatial unit. Its basic properties such as size, shape and location are determined mainly by the sensor (e.g. its altitude, look angle, field of view etc.) and not directly by the properties of the ground. Commonly, the area represented by a pixel will contain more than one thematic class as shown in Figure 7-23. This situation may arise for many reasons and makes the data unsuitable for use with some conventional image analysis techniques. Consequently, the use of conventional 'hard' (one pixel one class) approaches to mapping from remotely sensed data will result in error if mixed pixels are present. The magnitude of the problem is also clearly a function of the proportion of mixed pixels within the remotely sensed data set (Foody, 2004).

There are many different methods of implementing sub-pixel classification in remote sensing. Two common approaches include linear mixing model and fuzzy set classification will be here reviewed.

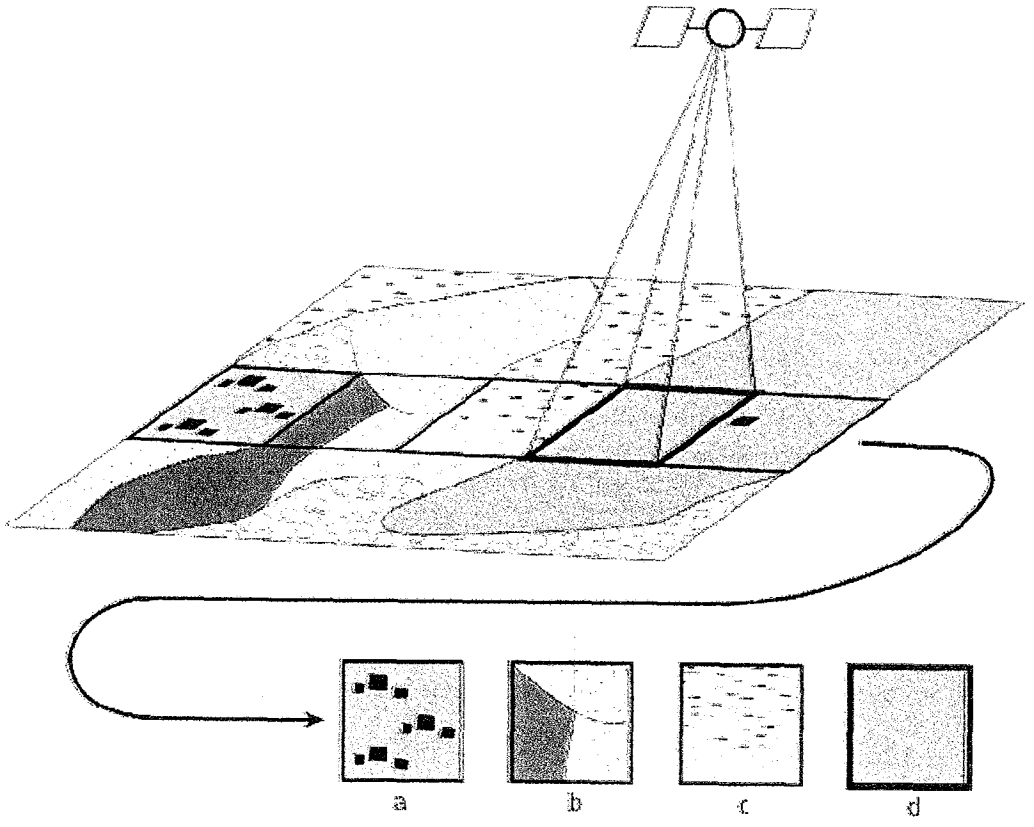


Figure 7-23: Some common origins of mixed pixel problems (From Foody, 2004).

A. Linear Mixing Model

The idealized, pure signature for a class is called an endmember. Because of sensor noise and within-class signature variability, endmembers only exist as a conceptual convenience and as idealizations in real image (Showengerdt, 1997). The spatial mixing of object within a ground projected field of view (GIFOV) and the resulting linear mixing of their spectral signature are illustrated in Figure 7-24. Usually, linear mixing modeling assumes a single reflectance occurs within the GIFOV.

If these endmembers are indeed "pure," and if they are an exhaustive basis for all spectral in the image, the spectral vector for any pixel must lie within the convex hull defined by the enveloped around the endmember. As shown in Figure 7-25, there are three endmembers in two-dimensional feature space. The class fractions determine the location of the mixed-pixel vector in feature space. The inversion problem, termed unmixing, is required to estimate the fractions of each for a given pixel vector (Schowengerdt, 1997).

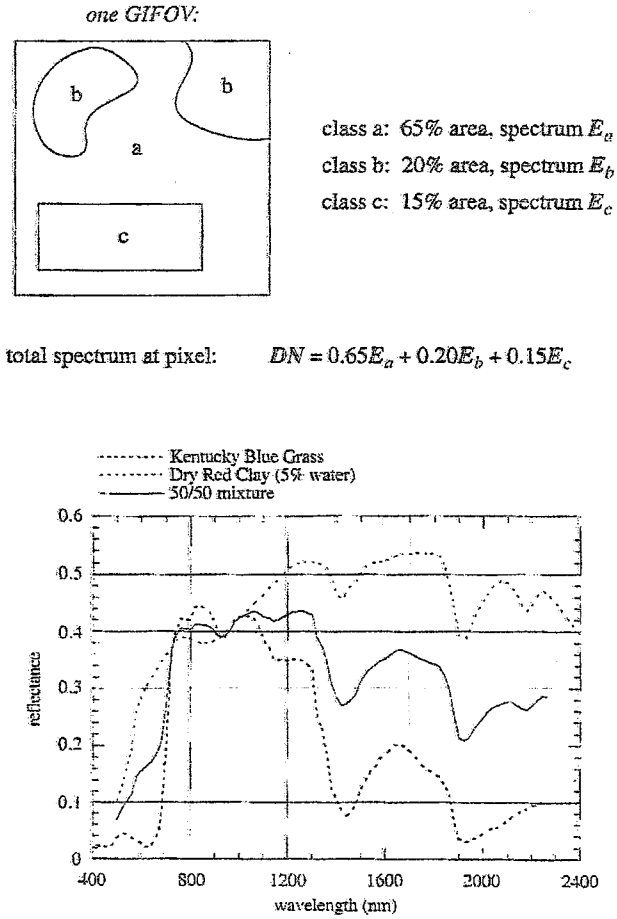


Figure 7-24: The linear mixing model for a single GIFOV. (From Schowengerdt, 1997)

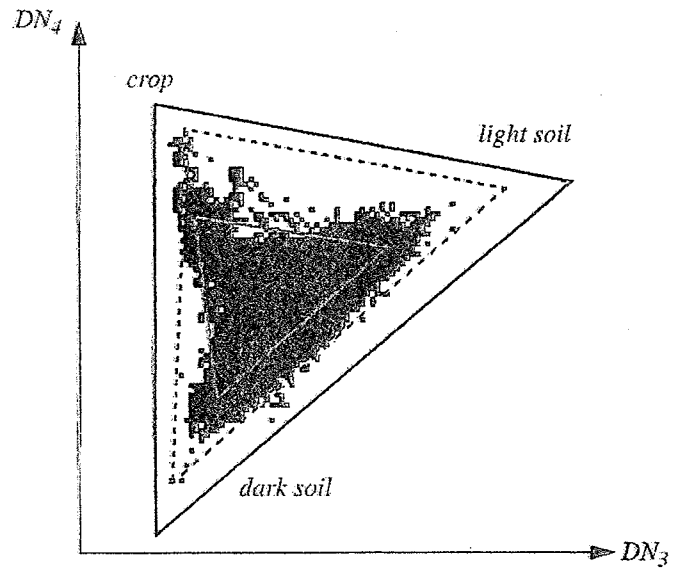


Figure 7-25: Three possible choices for endmembers for the classes "dark soil," "light soil," and " crop." (From Schowengerdt, 1997)

The basic and widely used linear mixture model is based on the assumption that the spectral response x_k of a pixel k is a linear weighted sum of spectral responses of its component classes (Foody, 2004), expressed by

$$x_k = E\int + \varepsilon \quad (7\text{-}12)$$

in which E is a q by c matrix in which q represents the number of wavebands and c the number of classes, the term f represents a vector, of length c, that expresses the proportional coverage of the classes in the area represented by the pixel and ε is a term that represents the residual error. The columns of the matrix E are end-member spectra, essentially the characteristic spectral responses of the classes. These can be difficult to define appropriately and are commonly determined from 'pure' training areas. Once defined, the mixture model may be used to estimate the class composition of a pixel, represented by f, from its spectral response subject to the constraints that

$$\int i \geq 0, i = 1,... c \quad (7\text{-}13)$$

and

$$\sum_{n=1}^{c} \int_{i} = 1 \quad (7\text{-}14)$$

This simple approach to deriving sub-pixel information has been used widely in remote sensing. Commonly, the spatial distribution of the land cover classes is represented through fraction images. Typically there is one fraction image for each class and in this representation the pixel's grey level scale is directly related to the proportional coverage of the class in the area represented by the pixel as shown in Figure 7-26.

Foody (2004) stated that despite its simplicity and ability to derive accurate sub-pixel estimates under certain circumstances there are some concerns about the use of this approach to pixel unmixing. The use of the least squares error criterion, for example, makes the unmixing analysis prone to problems associated with outliers. A simple alternative that reduces this effect is to base the unmixing analysis on a least median square criterion In addition, this approach requires that the number of end-members is less than the number of spectral bands used in the analysis.

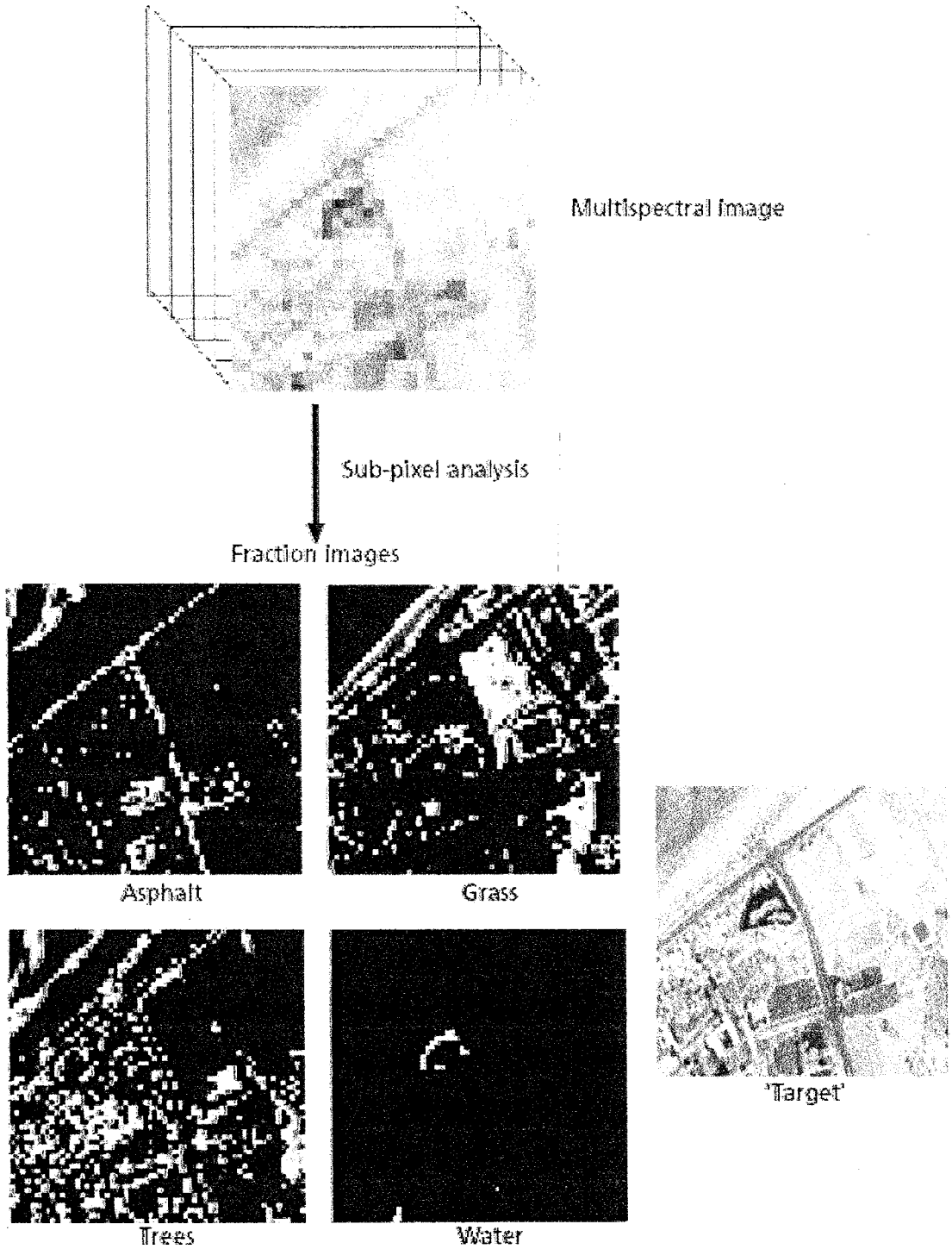


Figure 7-26: An example of a typical sub-pixel analysis (From Foody, 2004)

B. Soft (Fuzzy) Classification

In soft thematic mapping, each pixel is not forced or constrained to show membership to a single class. Consequently, each pixel may display multiple and partial class membership; since a single class label is no longer forced upon pixels the output is fuzzy (Foody, 2004).

Jensen (2005) illustrated the use of fuzzy classification logic to discriminate among the three hypothetical land covers as shown in Figure 7-27. The vertical boundary of water at brightness value 30 is placed by a graded boundary that represented a gradual transition from water to forested wetland. In the language of fuzzy set theory, *BVs* of less than 24 have a *membership grade* of 1 for water, and those greater than about 70 have a membership grade of 1 for upland forest. At several other locations a BV may have a membership grade in two class. For example, at BV 30 we have membership grades of 0.5 water and 0.5 of forested wetland. At BV 60 the membership grades are 0.7 for forested wetland and 0.3 for upland forest. This membership grade information may be used by the analyst to create a variety of classification map.

Foody (2004) stated that frequently, the grade or strength of membership a pixel displays to each class is derived and with many of the popular approaches to soft classification the measure(s) of the strength of class membership derived for a pixel are taken to reflect the relative proportion of the classes in the area represented by the pixel. Many methods may be used to derive a measure of the strength of class membership that can be used to form a soft thematic map included Maximum likelihood classification, Fuzzy c-mean (FCM) and Possibilistic c-Means (PCM) and Neural Networks. The briefly characteristics of these are summarized based on Foody (2004) as following.

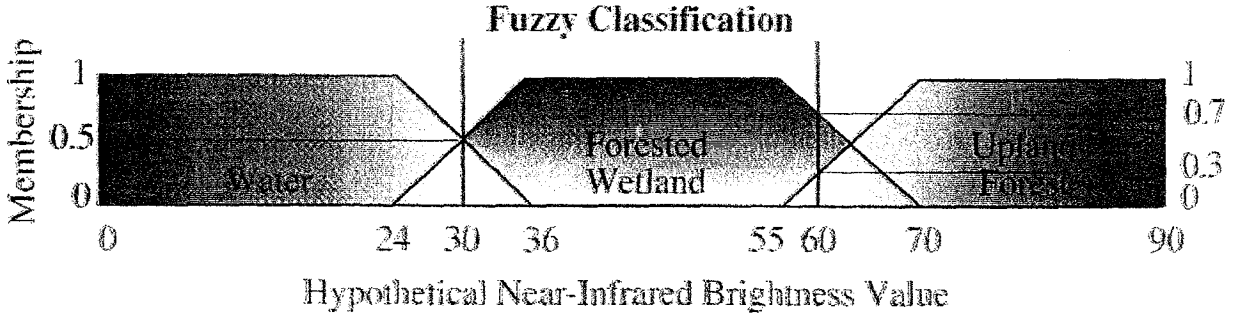


Figure 7-27: The logic of fuzzy classification (From Jensen, 2005).

A Maximum likelihood classification

Although the maximum likelihood classifier is a conventional statistical classification technique that allocates each pixel to the class with which it has the highest likelihood or a posterior probability of membership it can be softened. The basis of the maximum likelihood classification, and other related probabilistic classifiers which may be used in the same way, is the probability density function, which may be derived from,

$$p\left(x_k|i\right)=\frac{1}{\sqrt{2\pi}\sqrt{\left|M_i\right|}}exp\left(-\frac{1}{2}D^2\right) \quad (7\text{-}15)$$

where $p(x_k|i)$ represents the probability density function for the pixel *k* with the data vector x_k as a member of class i, M_i is the variance – covariance matrix for class i and D^2 is the Mahalanobis distance between the pixel *k* and the centroid of class i. The Mahalanobis distance may be calculated from,

$$D^2=\left(x_k-v_i\right)^T M_i^{-1}\left(x_k-v_i\right) \quad (7\text{-}16)$$

where v_i is the mean vector for class i. The calculated Mahalanobis distance may be converted to a typicality probability by reference to a chi-squared distribution. These probabilities lie on a 0-1 scale and represent the probability of observing a Mahalanobis distance as extreme as that observed for a particular pixel with respect to the specified class. Thus, the typicality probability provides an indication of the closeness of the pixel to a single specified class centroid. That is, the typicality provides a measure of the absolute strength of membership a pixel has to an individual class, irrespective of its membership to any other class. Note that the sum of the typicality to all specified classes needs not sum to 1.0 for any one pixel. These properties of the typicality make it a very different measure of class membership to the posterior probability.

In a standard maximum likelihood classification, each pixel is allocated to the class with which it has the highest posterior probability of class membership, which is calculated from,

$$L\left(i|x_k\right)=\frac{P_i p\left(x_k|i\right)}{\sum_{j=1}^{c}P_j p\left(x_k|j\right)} \quad (7\text{-}17)$$

where L(i|xk) represents the posterior probability of pixel with the data vector xk belonging to class i, Pi the a priori probability for class i, and c the total number of classes. Posterior probabilities of class membership lie on a 0-1 scale but they are also constrained to sum to 1.0 for each pixel and are derived with respect to all defined classes. Thus, the posterior probability provides a relative

measure of class membership. This situation makes the posterior probability attractive as an indicator of sub-pixel proportions. In practice, useful sub-pixel thematic information may be derived by this approach.

B. Fuzzy c-means (FCM)

The FCM is a fuzzy classifier that has been widely used for the derivation of sub-pixel scale thematic information. The FCM is, essentially, a non-hierarchical clustering algorithm that may be used to sub-divide a data set into c classes. Pixels are initially assigned randomly to classes before being moved iteratively to other classes with the aim of minimizing the generalized least-squared error,

$$J_m(U,v)=\sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^m(d_{ik})^2 \quad (7\text{-}18)$$

where U is a fuzzy c-partition of the data set containing n pixels (x_1, x_2, ..., x_n), v_i is the centre of cluster i, d_{ik} is the distance between x_k and v_i measured using an appropriate Sub-Pixel Methods in weight matrix and m is a user defi ned weighting component that lies within the range $1 \le m \le \infty$ which determines the degree of fuzziness of an analysis. When m=1 a conventional hard classification is obtained in which each pixel is associated unequivocally with just one class. For the derivation of sub-pixel scale information, it is, therefore, important that m>1, ensuring that multiple and partial class membership is allowed in the output. There is no optimal value of m and most studies have used a value in the range 1.5<m<3.0. In a soft classification, attention is focused on the elements of U, u_{ik}, that represent the grade of membership of a case to a class. These membership values are constrained to satisfy,

$$u_{ik} \in [0,1] \quad (7\text{-}19)$$

$$\sum_{k=1}^{n} u_{ik} > 0, i = 1,\ldots,c \quad (7\text{-}20)$$

$$\sum_{k=1}^{c} u_{ik} = 1, k = 1,\ldots,n \quad (7\text{-}21)$$

In a fuzzy c-partition of a data set, the membership functions characterize the membership of each case in all classes. Memberships close to unity indicate a high degree of similarity between a case and a class whereas memberships close to zero indicate little similarity between a case and a class. It is apparent from the constraints defined above that the memberships lie on a 0-1 scale and sum to unity for each pixel. These constraints make the fuzzy memberships appear similar in general terms to posterior probabilities, although they are in reality quite different. The memberships are calculated from,

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}}} \quad (7\text{-}22)$$

(Bezdek et al., 1984; Schowengerdt, 1997).

C. Possibilistic c-Means (PCM)

The PCM is the possibilistic counterpart of the FCM. The main difference between the PCM and the more widely used FCM is the removal of the constraint for the memberships to sum to one for each pixel in the PCM. With the PCM the strength of membership is derived from,

$$u_{ik} = \frac{1}{1 + \left(\frac{d_{ik}^2}{\eta_i} \right)} \quad (7\text{-}23)$$

where η_i is a parameter that specifies the distance at which the membership to a class equals 0.5. The main attraction of the PCM for the derivation of sub-pixel scale thematic information is that, like the typicality probability, the membership values derived are measures of the absolute strength of class membership. Consequently, the memberships derived from the PCM are not affected by the presence of untrained classes. Although more accurate predictions of sub-pixel class composition may be derived from the FCM than PCM, the presence of untrained classes can markedly degrade the accuracy of sub-pixel estimates from the FCM while not affecting those from the PCM. Thus, in situations when the analyst may believe that the set of classes contained within the imaged area has not been defined exhaustively it may be preferable to use the PCM rather than FCM for the derivation of sub-pixel scale thematic information.

D. Neural networks

Neural networks have recently become a very popular tool for the derivation of sub-pixel scale thematic information from remotely sensed data. A variety of types of network have been used for the derivation of sub-pixel scale thematic information, notably feedforward networks.

Feedforward networks have been widely used for the classification of remotely sensed data. A variety of feedforward networks have been used for estimating sub-pixel scale thematic information, including the multi-layer perceptron, radial basis function and probabilistic neural networks. One attractive feature of feedforward networks for soft classification is the ability to directly include mixed pixels in the training stage, as the desired output vector for each training case can be specified; training may, however, be based upon pure pixels only. Once trained a

feedforward neural network may be used to predict the class membership properties for other pixels in the data set. The magnitude of the outputs units used to indicate the strength of class membership, which may be used to indicate sub-pixel class proportional cover. The outputs of the different types of network should, however, be treated differently. While the activation level of output units from each type of network can be expressed on a 0-1 scale such outputs should not necessarily be viewed as proportions or probabilities. For example, the sum of the activation level of all output units for a multi-layer perceptron may differ markedly from 1.0 and of the networks identified above only the probabilistic neural network outputs values that are essentially posterior probabilities. However, in many studies the magnitude of the activation level of output units has been found to be strongly related to the class composition.

7.5.2.6 Artificial Neural Network (ANN) Classifier

Schowengerdt (1979 mentioned that the ANN algorithm is a recently popular nonparametric approach to classification. The decision boundaries by ANN are not fixed by a deterministic rule applied to the prototype training signature, but are determined in an iterative fashion by minimizing an error criterion on the labeling of the training data. In that sense, ANNs are similar to clustering algorithms.

ANNs simulate the thinking process of human beings, whose brains use interconnected neurons to process incoming information. A neural network reaches a solution not via a step-by-step algorithm or a complex logical program, but in a non-algorithmic, unstructured fashion based on the adjustment of the weights, connecting the neurons in the network (Jensen, 2005).

Keiner (1999) stated that ANN can be used in for a variety of remote sensing application, including classification, noise reduction, feature tracking, forecasting, and function approximation. Jensen (2005) mentioned that ANNs have been used to classify various types of remote sensor data and have in certain instances produces results superior to those of traditional statistics methods. This success can be attributed to two of the important advantages of neural networks:

- freedom from normal distribution requirements and
- ability to adaptively simulate complex and nonlinear pattern given proper topological structures.

Component and Characteristics of ANN

A neural network consists of a system of interconnected layers, which modify data as it passes through the network. The first layer of nodes (the input layer) brings the information to be processed into the network. Nodes in subsequent layers are called neuron, as they perform neuron-like functions. Eventually, the final produce the output of the network Keiner (1999).

The artificial neural network normally contains neurons arranged in three types of layers (1) an input layer, (2) a hidden layers, and (3) an output layer as shown in Figure 7-28. The hidden layer and the output layer contain processing elements at each node. The input layer nodes, on the other hand, are simply an interface to the input data and do not do any processing. The input patterns are the features used for classification. In the simplest case, they are the multispectral vectors of the training pixels, one band per node. Other features, such as a spatial neighborhood of pixels or multi-temporal spectral vectors, can also be used (Schowengerdt, 1997).

Jensen (2005) mentioned that ANN requires training and testing (classification) to extract useful information from the remotely sensed and ancillary data like a supervised classification. The neuron in the input layers might be the multispectral reflectance values plus their texture, surface roughness, terrain elevation, slope, aspect, etc. The use of neurons in the hidden layers enables the simulation of non-linear patterns in the input data. A neuron in the output layers might represent a single thematic map land-cover class, e.g., agriculture.

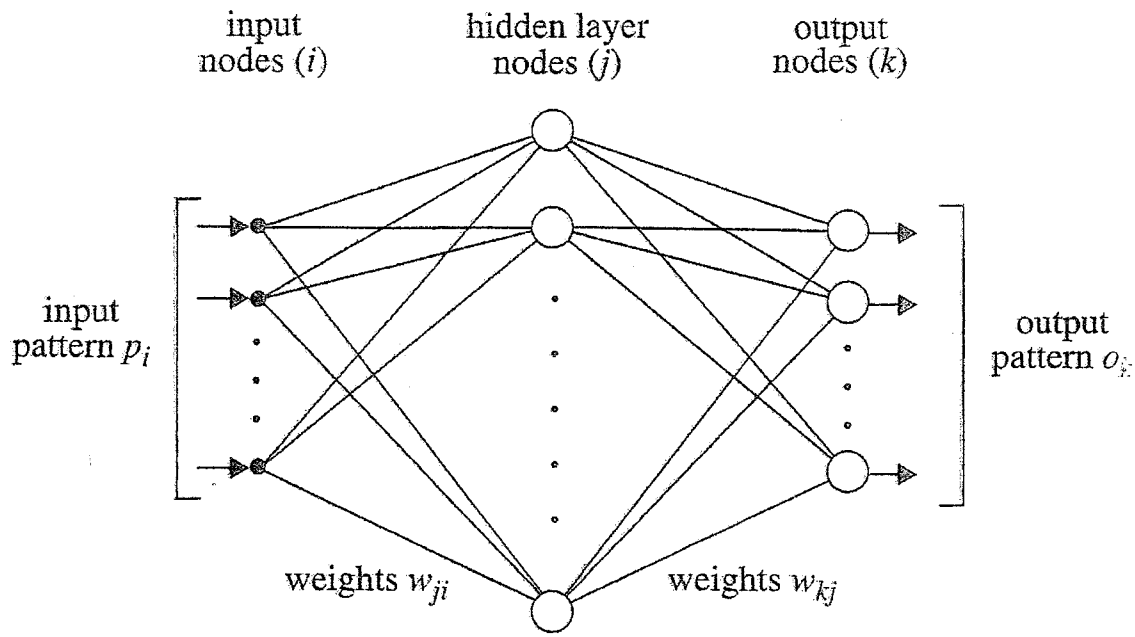


Figure 7-28: Basic structure of a three layer Artificial Neural Network (From Schowengerdt, 1997).

Within each processing node, there is a summation and transformation (Figure 7-29). At each hidden layer node, j, the following operation is performed on the input pattern, pi, producing output, hj, (Schowengerdt, 1997):

$$\text{Hidden layer:} \quad S_j = \sum_i w_{ji} p_i \text{ and } h_j = f(S_j) \tag{7-24}$$

Which is directed to each output layer node, k, where the output, ok, is calculated,

$$\text{Output layer:} \quad S_k = \sum_j w_{kj} h_j \text{ and } o_k = f(S_k) \tag{7-25}$$

The most widely-used transformation function is the sigmoid function (Figure 7-30):

$$f(S) = \frac{1}{1+e^{-s}} \tag{7-26}$$

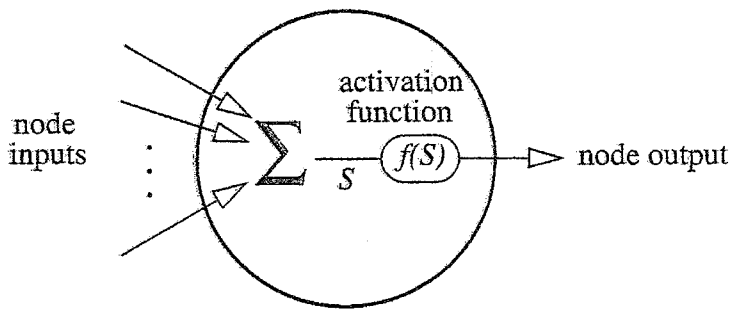


Figure 7-29: The components of a processing element (From Schowengerdt, 1997).

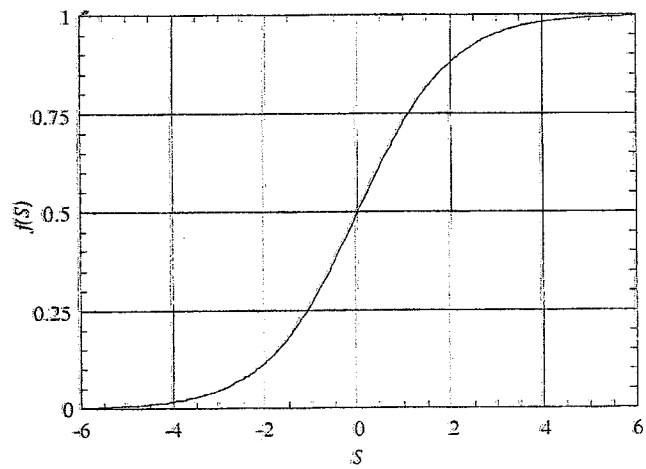


Figure 7-30: The sigmoid activation function (From Schowengerdt, 1997).

In addition, Jensen (2005) stated that ANN is defined by neurons, topological structure, and learning rules. The neuron is the fundamental processing unit of an ANN for computation. Each input x_i is multiplied by the scalar weight w_i to form $x_i w_i$, a term that is sent to the "summing unit" of the processing unit (Figure 7-31). An offset, b, may be added to the total. The summation output:

$$net = \sum_{i=1}^{p} (x_i w_i) + b \quad (7\text{-}27)$$

It referred to as net input, goes into an activation transfer function f that produces scaled neuron output y (between 0 and 1, or -1 to 1) through a transform algorithm. Thus, y can be calculated as

$$y = f(net) = f\left[\sum_{i=1}^{p} (x_i w_i) + b\right] \quad (7\text{-}28)$$

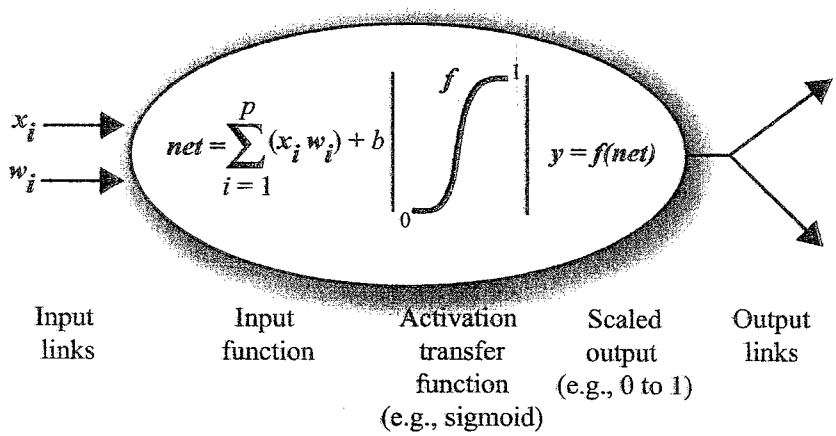


Figure 7-31: Mathematic model of neuron (From Jensen, 2005).

In addition, McCloy (2006) simply explained ANN algorithm with example as follows:

Consider that there are n inputs, representing the n input channels in the data, Xi,j = [x1, x2, x3, ..., xn]. Consider an ANN with one hidden layer with assigned weight W1 = [w1,1, w1,2, w1,3, ..., w1,n]. The activation (a) is then computed as:

$$a = \sum_{i=1}^{n} w_{1,i} \times x_i \quad (7\text{-}29)$$

The output is then found by thresholding the activation, a, by the use of a thresholding function. A simple thresholding function is the binary step function:

y = 1 if a >= threshold

y = 0 if a < threshold

In practice, the logistic function is a much more common thresholding function. The logistic thresholding sends most values to the extremes, but it provides a softening of the transformation in comparison with the binary step function (Figure 7-32):

$$y = 1/(1-e^{-a}) \quad (7\text{-}30)$$

Consider a simple case of two input values (x1, x2) that can take the value of 0 and 1. Initially set the two weights to unity and the threshold to 1.5. The result of the computation of the activation and the assignment of value to y are shown in Table 7-8. If the data space for the data in Table 7-8 is drawn as shown in Figure 7-33, then it can be seen that the threshold gives a linear decision surface. In this simple case the sum of the products at the decision surface will have the equation:

$$x1 \times w1 + x2 \times w2 = t \quad (7\text{-}31)$$

Equation 7-31 is a linear equation. Thus, this type of ANN is called a linear classifier.

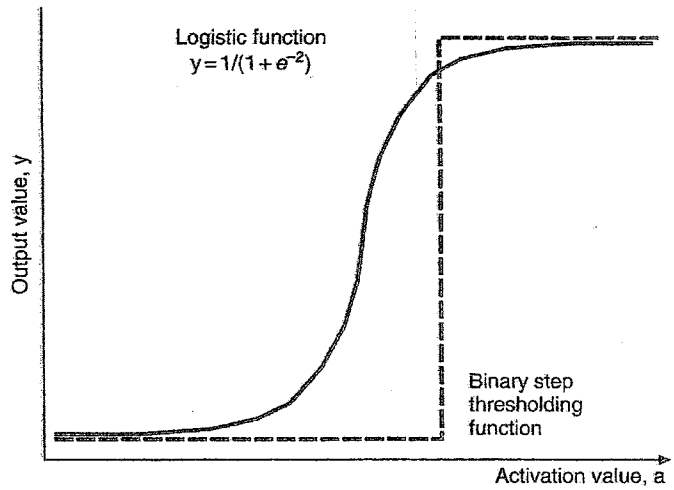


Figure 7-32: Two thresholding functions used to transform the activation to output values at a synapse (From McCloy, 2006).

Table 7-8: The input values, activation and output value for simple two-channel ANN

x_1	x_2	Activation, a	Output, y
0	0	0	0
0	1	1	0
1	0	1	0
1	1	2	1

From McCloy (2006

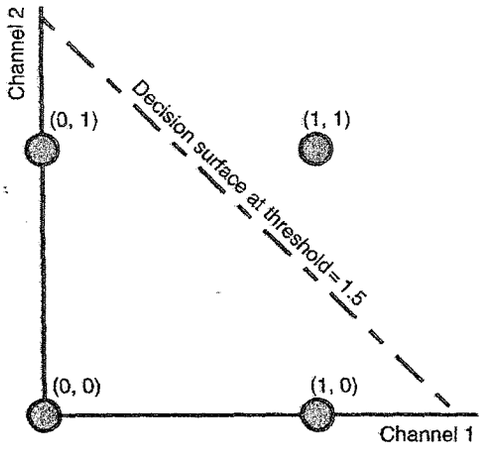


Figure 7-33: The data space and the decision surface for a simple two channel, single hidden layer ANN (From McCloy, 2006).

In general, there are two commonly used ANN learning patterns (Jensen, 2005):

1. a pattern associator based on a least-mean-squared algorithms
2. an *error back-propagation algorithm.*

The Pattern Associator ANN Algorithm

A pattern associator (PA) model is a primitive neural network that consists of only one input layer and one output layer and one output layer and operates under the assumption of linearity It based on the least-mean-square (LMS) algorithm. The LMS algorithm is driven by the difference between the target activation and the obtained activation. The idea is to adjust the weights of the connection to reduce the mean-square error (Jensen, 2005).

First, the network is initialized by setting all the weight to zero:

$$wk(1) = 0 \text{ for } k = 1,2,3,\ldots p \qquad (7\text{-}32)$$

where:

p is the number of inputs

wk(1) is the weight associated with input k in the first iteration.

For every iteration t, t = 1,2,..., the obtained activation y can be calculated:

$$y(t) = \sum_{i=1}^{p} \left[x_i(t) \hat{w}_i(t) \right] \qquad (7\text{-}33)$$

where:

xj(t) is the value of input j at iteration t.

Thus the difference between the desired activation d(t) and obtained activation y(t) at iteration t can be computed:

$$e(t) = d(t) - y(t) \qquad (7\text{-}34)$$

Then, the difference between the desired activation d(t) and obtained activation y(t) at iteration t can be computed:

$$\hat{w}_k(t+1) = w_k(t) + \eta \cdot e(t) \cdot x_k(t) \quad (7\text{-}35)$$

where:

η is a positive constant known as the learning rate.

The traditional PA usually uses a linear transfer function in its output neuron, so that the activation of the output neuron is simply equal to the net input. In this sense, the LMS learning process emulates a linear regression analysis, except that it is free from any assumption of data distribution. The traditional PA fails to achieve a good result when the data being modeled are not linear. Variants to the linear PA adopt nonlinear transform functions such as a continuous sigmoidal function that resembles regression with a logarithmic transformation. These variants have the ability of modeling simple nonlinearity without distribution constraints.

The Error Back-propagation ANN Algorithm

Back-propagation (BP) neural networks were first proposed by Werbos (1974) to improve the efficiency of multilayer network training (Jensen et al., 1999). This network has a hierarchical architecture with fully interconnected layers. Typically, one or more hidden layers are used to enable the network to learn complex tasks. The training is based on an error-correction learning rule. Each neuron in the network may include a nonlinear transfer function at the output end, producing smooth signals to other neurons (Jensen, 2005). One of the common forms of nonlinearity is a sigmoidal transfer function defined by the logistic function:

$$f(net) = \frac{1}{1+e^{(-net)}} \quad (7\text{-}36)$$

Initialized with all the synaptic weights and thresholds set to small random numbers, the network is fed with training patterns. Each learning iteration of the network consists of two passes: a forward pass and a backward pass. In the forward pass, the net input of the jth neuron of layer l is calculated:

$$net_j^{\,l}(t) = \sum_{i=0}^{p} (y_i^{\,(l-1)}(t) w_{ji}^{\,l}(t)) \quad (7\text{-}37)$$

And the output activation is computed using the transfer function (Eq. 7-x):

$$y_j^{\,l}(t) = f\left[net_j^{\,l}(t)\right] \quad (7\text{-}38)$$

The difference between the desired response and output activation of neuron j at the output layer is obtained by:

$$e_j(t) = d_j(t) - o_j(t) \tag{7-39}$$

In the backward pass, a local error gradient, δ, is computed layer by layer:

$$\delta_j^{\ L}(t) = e_j^{\ L}(t) o_j(t)[1 - o_j x(t)] \tag{7-40}$$

for neuron j in output layer L, and

$$\delta_j^{\ l}(t) = y_j^{\ l}(t)[1 - y_j(t)] \sum_k \delta_j^{\ (l+1)}(t) w_{kj}^{\ (l+1)}(t) \tag{7-41}$$

for neuron j in hidden layer L.

With this information, the synaptic weights in layer l can be adjusted according to a generalized LSM rule:

$$w_{ji}^{\ l}(t+1) = w_{ji}^{\ l}(t) + \alpha[w_{ji}^{\ l}(t) - w_{ji}^{\ l}(t-1)] + \eta \delta^l{}_j(t) y_i^{\ (l-1)}(t) \tag{7-42}$$

where η is the learning rate and α is the momentum constant for speeding up learning without running into the risk of oscillation.

In special case when α equal zero, the generalized LSM rule becomes the standard LSM. In the backward pass, weights are adjusted in a manner that changes are proportional to the error gradient. Because this pass is started from the output layer and error correction is propagated backward to the precious layer, this process is called error back propagated. The knowledge inherent in the training data is obtained through the iteration of forward and backward passed until all parameters in the network are stabilized (Jensen, 2005).

Advantage and disadvantage of ANN were summarized by Jensen (2005) as in Table 7-9.

Table 7-9: Advantage and disadvantage of ANN algorithm.

Advantage	Disadvantage
• A single neuron simulates the computation of a multivariate linear regression model. • A neural network makes no a priori assumptions of normal and linear data distribution due to its operation in a nonparametric fashion. • Neural networks are able to learn from existing examples adaptive, which makes the classification objective. • The nonlinear patterns are "learned" from the empirical examples instead of pre-specified" by an analysis based on prior knowledge of the datasets. • The noisy information inevitably included in the examples supplied a trained neural network with the ability to generalized, which makes neural networks robust solutions in the presence of previous unseen, incomplete, or imprecise data. • A neural network can embrace data in all formats as long as the data are converted to a numeric representation. • Neural network are tolerant of noise and missing data and attempt to find the best fit for input patterns. • Neural networks continuously adjust the weights as more training data are provided in a changing environment.	• Despite the excellent performance of neural networks in image classification, it is usually difficult to explain in a comprehensive fashion the process through which a given decision or output has been obtained from a neural network. The rules of image classification and interpretation learned by the neural network are buried in the weights of the neurons of the hidden layers. It is difficult to interpret these weights due to their complex nature. A neural network is often accused of being a black box. • Using neural network, an analyst might find it difficult to gain an understanding of the problem at hand because of the lack of explanatory capability to provide insight into the characteristics of the dataset. • It is difficult to incorporate human expertise to simplify, accelerate, or improve the performance of image classification; a neural network always has to learn from scratch.

Form Jensen, 2005.

7.5.3 Nonmetric Classification

Duda et al. (2001) mentioned that nonmetric methods turned away from description patterns by vectors of real numbers and toward using lists of attributes. Nonmetric algorithms used nominal data for classification by rule-based or synthetic pattern recognition methods. The most commonly used nonmetric classification based on rule-based decision tree classifiers is expert system.

7.5.3.1 Expert Systems

Jensen (2005) defined a knowledge-based expert system as a system that used human knowledge to solve problems that normally would require human intelligence. It is the ability to solve problems efficiently and effectively in a narrow problem area and to perform at the level of an expert. Expert systems represent the expert's domain (i.e., subject matter) knowledge base as data and rule within the computer. The rules and data can be called upon when needed to solve the problems. A different problem within the domain of knowledge base can be solved using the same program without reprogramming. Examples of the use of expert system in remotely sensed data classification were described by Jensen (2005) and here summarized in Table 7-10.

Table 7-10: Examples of the use of expert system in remotely sensed data classification

Applications	**Reference**
Overview of expert systems related to image processing	Matsuyama (1989)
Inventory land use cover by using texture and expert system	Moller-Jensen (1997)
Extract land-cover information from MODIS by using a decision tree classifiers	Munchoney et al. (2000)
Summary of numerous applications of hierarchical decision-tree classifiers	Tso and Mather (2001)
Assessed the effectiveness of decision-tree methods for land-cover classification.	Pal and Mather (2003)
Apply an expert system to model what it would take to stabilize the sea water level to the 1960 level within 10 to 20 years	Krapivin and Phillip (2001)
Change detection in complex heterogeneous environments	Stefanov et al., 2001;Yang and Chung, 2002; Stow et al., 2003

Modified from Jensen (2005)

A Component of a Knowledge-Based Expert System

A knowledge-based expert system consists of the component as shown in Figure 7-34 including (1) Human expert, (2) User interface (3) Knowledge base (rule-based domain) (4) Inference engine (5) On-line databases, and (6) Users.

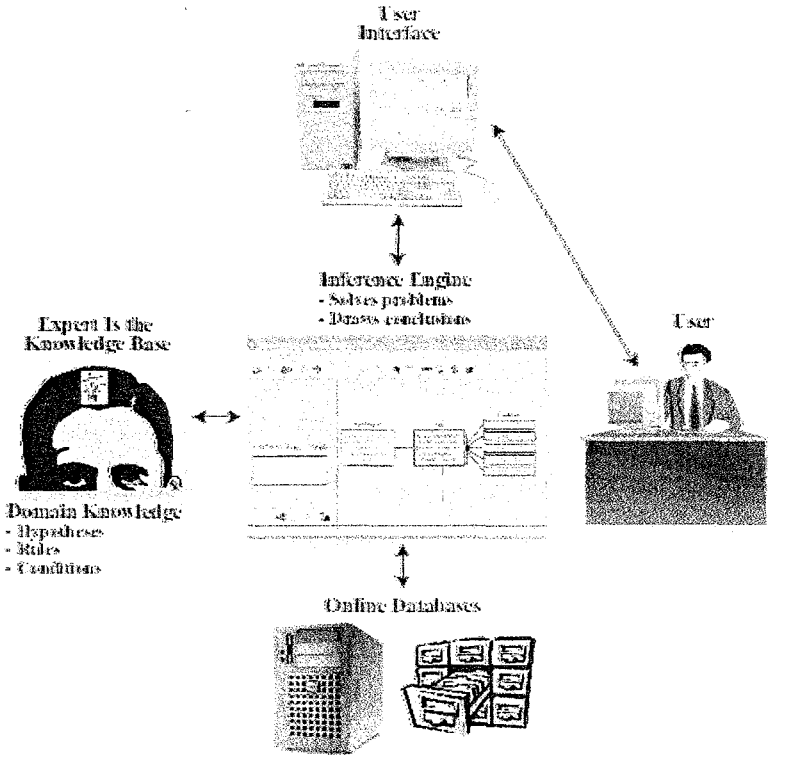


Figure 7-34: The Components of a Typical Rule-based Expert System (From Jensen, 2005).

Characteristics of major components of expert systems were summarized based on Jensen (2005) in following section:

A.1 Expert System User Interface

The expert system user interface should be easy to use, interactive, and interesting. It should also be intelligent and accumulate user preference in an attempt to provide the most pleasing communication environment possible. This expert system shell was built using object-oriented programming and is easy to use. All of the hypotheses, rules, and conditions for an entire expert system may be viewed and queried from the single user interface.

A2. Creating the knowledge base

There are three different problem-solving approaches using by some laypersons and scientists to process their knowledge:

1. Algorithms using conventional computer programs;
2. Heuristic knowledge-based expert systems, and
3. Artificial neural networks.

Algorithms using conventional computer programs

Conventional algorithmic computer programs contain little knowledge other than the basic algorithm for solving a specific problem, the necessary boundary conditions, and data. The knowledge is usually embedded in the programming code. As new knowledge becomes available, the program has to be changed and recompiled.

Heuristic Knowledge-based Expert Systems

Knowledge-based expert systems collect many small fragments of human know-how for a specific application area (domain) and place them in a knowledge base that is used to reason through a problem, using the knowledge that is most appropriate. Heuristic knowledge is defined as involving or serving as an aid to learning, discovery, or problem solving by experimental and especially by trial-and-error methods. Heuristic computer programs often utilize exploratory problem-solving and self-educating techniques as the evaluation of feedback to improve performance (Merriam-Webster, 2003).

Artificial Neural Network

Artificial Neural networks simulate the thinking process of human beings, whose brains use interconnected neurons to process incoming information. A neural network reaches a solution not via a step-by-step algorithm or a complex logical program, but in a non-algorithmic, unstructured fashion based on the adjustment of the weights, connecting the neurons in the network.

The knowledge representation process normally involves encoding information from verbal descriptions, rules of thumb, images, books, maps, charts, tables, graphs, equations, etc. The knowledge base contains sufficient high-quality rules to solve the problem under investigation. Jensen (2005) state that rules are normally expressed in the form of one or more "IF condition THEN action" statement. The condition portion of a rule statement is usually a fact. When certain rules are applied, various operations may take place such as adding a newly derived derivative fact to the database or firing another rule. Rules can be implicit (slope is high) or explicit (e.g., slope > 70%). It is possible to chain together rules, e.g., IF c THEN d; IF d THEN e, therefore IF c THEN e. It is also possible to attach confidence (e.g., 80% confident) to facts and rules. For example a typical remote sensing might be:

IF blue reflectance is (Condition) < 15%

AND green reflectance is (Condition) < 25%

AND red reflectance is (Condition) < 15%

AND NIR reflectance is (Condition) > 45%

THEN there is strong suggestive evidence (0.8) that the pixel is vegetated.

The best way to conceptualize an expert system is to use a decision-tree-structure where rules and conditions are evaluated in order to test hypotheses. When decision trees are organized with hypotheses, rules, and conditions, each hypothesis may be thought of as the trunk of a tree, each rule a limb of a tree, and each condition a leaf. This is commonly referred to as a hierarchical decision-tree classifier. The purpose of using a hierarchical structure for labeling objects is to gain a more comprehensive understanding of relationships among the objects at different scales of observation or at different level of detail.

McCloy (2006) stated that the concept behind hierarchical classifiers is that different criteria characterize the classes at the different hierarchical levels and thus different criteria should be used for discrimination of the classes at these different levels. In general, the hierarchy process from the most general to the most detail. If the hierarchical levels are set as in Table 7-11, then the criteria for establishing the classes at each level can be seen to be quite different, suggesting different classification strategies may be useful at the different levels.

ERDAS (2002) stated that the decision tree grows in depth when the hypothesis of one rule is referred to by a condition of another rule. The terminal hypotheses of the decision tree represent the final classes of interest. Intermediate hypotheses may also be flagged as being a class of interest. This may occur when there is an association between classes. Figure 7-35 represents a single branch of a decision tree depicting a hypothesis, its rule, and conditions. In this example, the rule, which is Gentle Southern Slope, determines the hypothesis, Good Location. The rule has four conditions depicted on the right side, all of which must be satisfied for the rule to be true. However, the rule may be split if either Southern or Gentle slope defines the Good Location hypothesis. While both conditions must still be true to fire a rule, only one rule must be true to satisfy the hypothesis (Figure 7-36).

Table 7-11: Hierarchical levels, classes and criteria

Hierarchical Level	Classes	Criteria	Comments
1	1-land 2-water 3-cloud	High NIR values Low NIR values Low TIR values	Use NIR and TIR to discriminate these classes
1.1	1-green vegetation 2-brown vegetation 3-soil/rock/artificial 4-shadow	High in Vegetation indices Low in VI's, high in NIR Low in VI's, high in NIR Low in VI's, low in NIR	Use VI's, NIR and other visible bands to discriminate these classes
1.2	1-open water 2-swamp and marsh	Low NIR, low VI's, uniform Low NIR, moderate, variable VI's	
1.3	1-none		
1.1.1	1-herbaceous 2-woody 3-desert	Seasonal phenology No phenology except deciduous No phenology	Temporal dynamics and VI's
1.1.2	1-none		
1.1.3	1-rough 2-smooth	Texture	Texture
1.1.4	1-none		
1.2.1	1-sediment load 2-chlorophyll load	Response value in the Blue-green part of the spectrum	End Member analysis

From McCloy (2006)

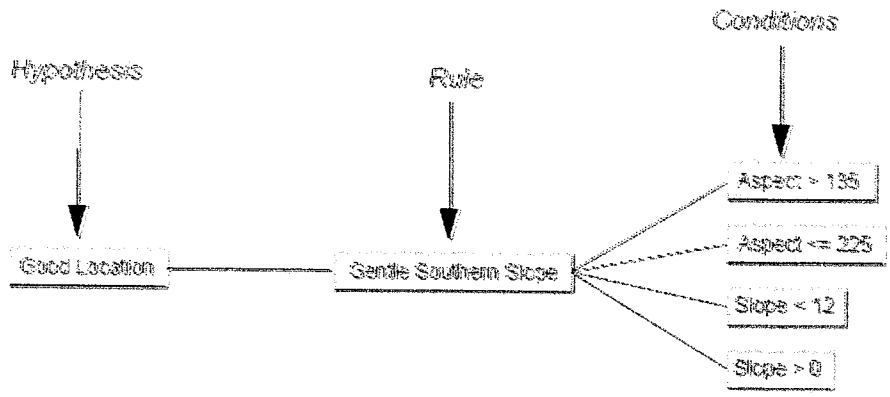


Figure 7-35: Example of a Decision Tree Branch (From ERDAS, 2002)

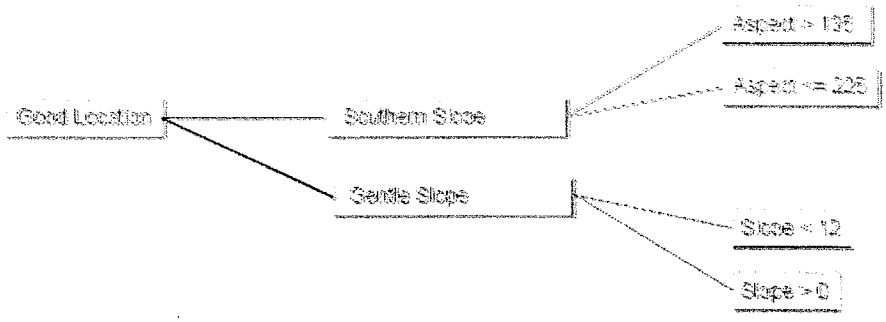


Figure 7-36: Split Rule Decision Tree Branch (From ERDAS, 2002)

A decision tree takes as input an object or situation described by a set of attributes and returns a decision. The input attribute can be discrete or continuous. The output values can also be discrete or continuous. Learning a discrete-valued function is called **classification learning**. Learning a continuous function is called **regression**. One of the operation is Boolean classification wherein each example classified as true (positive) or false (negative). A decision tree reaches its decision by performing a sequence of tests. Decision trees predict class membership by recursively partitioning a dataset into more homogeneous subsets (Jensen, 2005).

When building the knowledge domain, Jensen (2005) suggested that there were three steps involved:

Identify the hypotheses: The expert in charge of creating the knowledge domain identifies a hypothesis (problem) to be addressed. This may be a formal hypothesis to be tested using inductive logic and confidence levels or an informal hypothesis that is in search of a logical conclusion.

Specific the expert system rule: Heuristic rules that the expert has learned over time are the heart and soul of an expert system. Therefore, the creation of accurate, definitive rules is extremely important. Each rule provides the specific conditions to accept the hypothesis to which it belongs.

Specify the rule and condition: The expert would then specify one or more conditions that must be met for each rule.

An example of a human-derived decision tree expert system illustrated in Figure 7-37.

A.3 Inference Engine

Hypotheses, rules, and conditions are passed to the inference engine where the expert system is implemented. The inference engine interprets the rules in the knowledge base to draw conclusion. The inference engine may use backward- or forward-chaining strategies or both. Both backward and forward inference processes consist of a chain of steps that can be traced by the expert system. This enables expert systems to explain their reasoning processes.

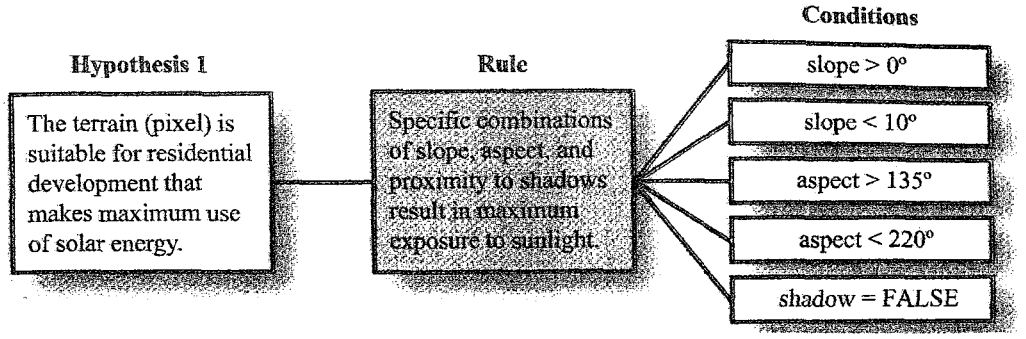


Figure 7-37: A human-derived decision-tree expert system with a rule and conditions (From Jensen, 2005).

A.4 On-line Databases

The rule and conditions may be applied and evaluated using data and/or information stored in on-line databases. The databases can take a variety of forms. It can be spatial and consist of remotely sensed images and thematic maps in raster and vector format. However, the database may also consist of charts, graphs, algorithms, pictures, and text that are considered important by the expert. The database should contain detailed, standardized metadata.

Jensen (2005) demonstrated that expert systems in remote sensing research may be classified white fir based on the use of formal rules developed by a human expert or derived automatically by an inductive machine-learning algorithm based on training data that is input by humans into system. The hypothesis (class), rules (variables) and conditions was summarized in Table 7-12. Also, Advantages and disadvantage of expert system were summarized in Table 7-13.

Table 7-12: A hypothesis (class), variables and conditions necessary to extract white fir (Abies concolor) forest cover information from Maple Mountain, Utah.

Hypotheses	Variables	Conditions
White Fir (Abies concolor)	Aspect	Aspect = 300 to 45 Degree
	Elevation	Elevation > 1200 m
	Slope	Slope = 25 to 50 Degree
	Multispectral	Remote sensing reflectance
		TM Band 1 Blue = 44 to 52
		TM Band 2 Green = 31 to 40
		TM Band 3 Red = 22 to 32
		TM Band 4 NIR = 30 to 86
		TM Band 5 Mid IR = 19 to 47
		NDVI = 0.2 to 0.7

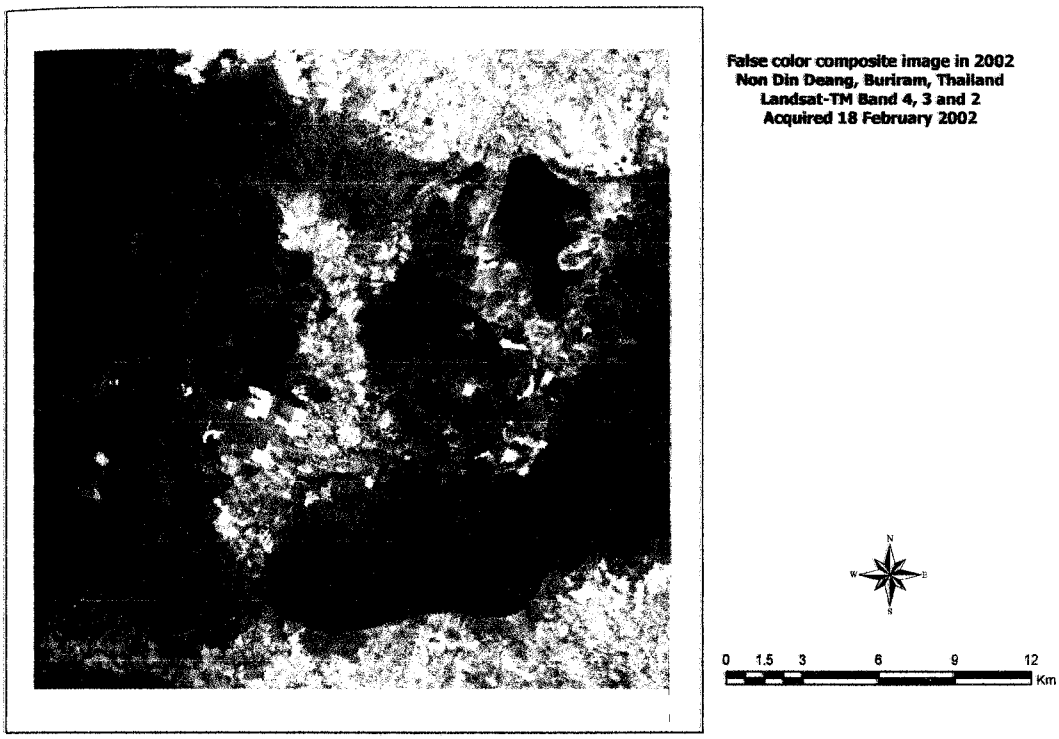
From Jensen, 2005.

Table 7-13: Advantages and disadvantage of expert system

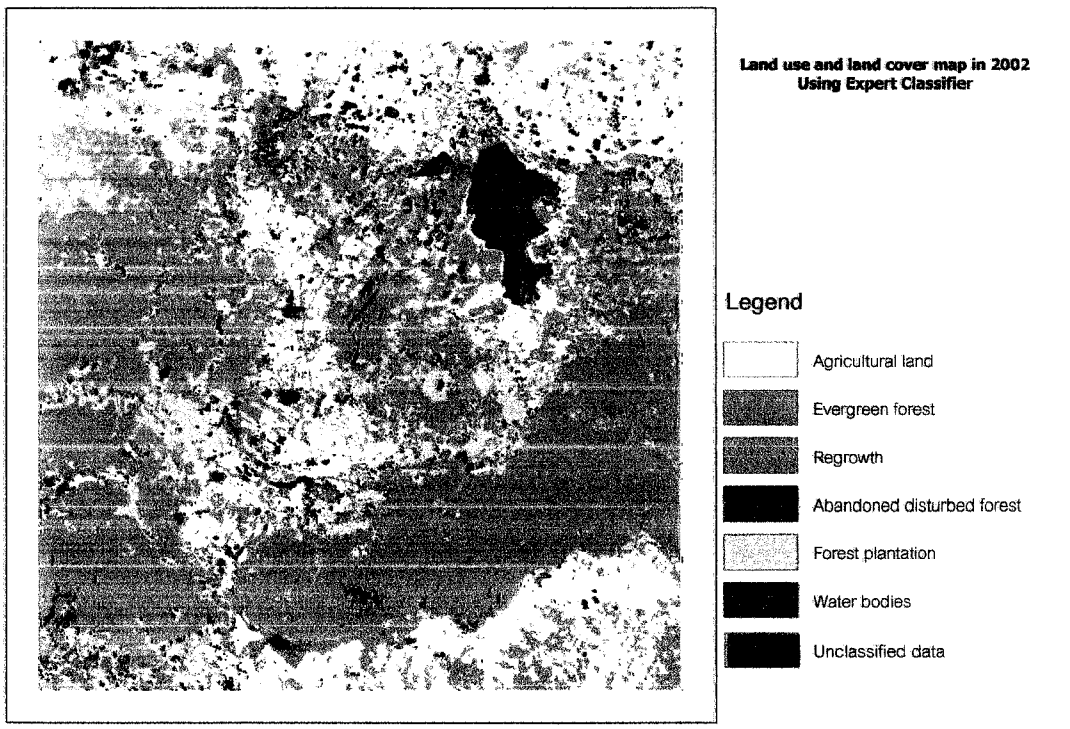
Advantage	Disadvantage
• Users can evaluate the output of the expert system and work backward to identify how a conclusion was reached. • Expert systems as nonmetric classification algorithm are being used such as decision trees, which make no assumption regarding the distribution of the data. • The decision tree can reveal nonlinear and hierarchical relationships among the input variables and use them to predict class membership. • A large body of evidence demonstrates the ability of machine-learning techniques (particularly decision trees and neural networks) to deal effectively with tasks that involve highly dimensional data.	• The knowledge in a traditional expert system that must be extracted from knowledgeable experts of a domain area may be subjective and incomplete. This is because the experts may have a biased or even incorrect understanding of reality, they may not be aware of underlying rules they have used, and they may have difficulty articulating these rules. • Knowledge in an expert system is represented by logical rules made up of binary predicates. Numerical attributes have to converted to binary true/false statements, which may cause a large amount of information to be lost in the simplification process. • Most rule-based expert systems fail to generalize a predictable inference if an appropriate match with the prefect rules that must be articulated by experts cannot be obtained.

From Jensen, 2005

Example of land use and land cover classification using expert classification was shown i Figure 7-38.



(a) False color composite data of Landsat-TM: Don Din Deang, Buriram, Thailand, Acquiring date: 18 February 2002



(b) Land use and land cover in 2002 map : Don Din Deang, Buriram, Thailand, Using Expert classification

Figure 7-38: Land use and land cover classification using expert classification

Chapter 8: Accuracy Assessment

Basically, the accuracy of maps from remotely sensed data is measured by two types of criteria: location accuracy and classification or thematic accuracy. Location accuracy refers to how precisely map items are located on the map relatively to their true location on the ground. Thematic accuracy refers to the accuracy of the map label describing a class or condition on the earth (Congalton and Green, 1999)

8.1 Source of Error in Remote Sensing-derived Thematic Products

Jensen (2005) categorized sources of errors in remote sensing-derived thematic products in subsequently groups as shown in Figure 8-1 included:

- Data Acquisition;
- Preprocessing;
- Information Extraction;
- Data Conversion;
- Error Assessment;
- Final Product Contains;
- Decision Making.

He briefly described these various as following:

A. Data Acquisition. First, error may be introduced during the remote sensing data acquisition process. It is quite common for remote sensing system detectors, cameras, RADARs, LIDARs, etc. to become miscalibrated. This can result in accurate remote sensing measurement (e.g., multispectral radiance, RADAR backscatter, LIDAR laser intensity). The aircraft or spacecraft platform may *roll, tilt,* or *yaw* during data acquisition. Erroneous commands may be put into the data-collection or inertial navigation system from ground control personnel. Finally, the scene being imaged may be randomly affected by unwanted haze, smog, fog, dust, high relative humidity, or sunglint that can dramatically affect the quality and accuracy of the information that is extracted.

B. Preprocessing. Considerable effort is spent preprocessing the remotely sensed data to remove both the *geometric and radiometric error* inherent in the original imagery. Unfortunately, even after preprocessing, there is always *residual geometric and radiometric error.* For example, residual geometric error may cause pixels still to be in their incorrect geometric locations. Similarly, even the best atmospheric correction will not result in a perfect relationship between percent reflectance measured on the ground and percent reflectance measured by an optical remote sensing system for the same geographic area. The residual geometric and radiometric error is passed to subsequent image processing functions.

C. Information Extraction. Sometime qualitative or quantitative information extraction techniques are based on flawed logic. For example, a hard land-cover classification scheme may not contain classes that are mutually exclusive, exhaustive, and hierarchical. Training sites may be incorrectly labeled during the training phase of a supervised classification. Clusters may be incorrectly labeled when performing an unsupervised classification. A human interpreted performing visual image interpretation may label a polygon incorrectly.

D. Data Conversion. Thematic information derived from the remotely sensed data may be in raster or polygon cartographic data structures. Sometimes it is desirable to convert the data from one data structure to another. The *conversion process* may introduce error.

E. Error Assessment. All above mentioned error accumulates in the remote sensing-derived information. Therefore, it is essential to conduct an error assessment (or accuracy assessment) to place confidence limits on the remote sensing-derived information. Unfortunately, the person conducting the error evaluation could make critical errors regarding the ground reference data sampling design, the number of samples data, and data autocorrelation. Hopefully, a robust multivariate statistical analysis is used so that confidence limits can be specified for the remote sensing-derived thematic information.

F. Final Product Contains. The final remote sensing-derived product is created when the thematic information is judged to be sufficiently accurate. Unfortunately, the analyst can still introduce error by not following standard cartographic procedures for the creation of scale bars and map legends. Metadata may also contain error.

G. Decision Making. Finally, the people who actually make the decision often do not understand or appreciate the amount of error that has accumulated in the remote sensing-derived thematic product.

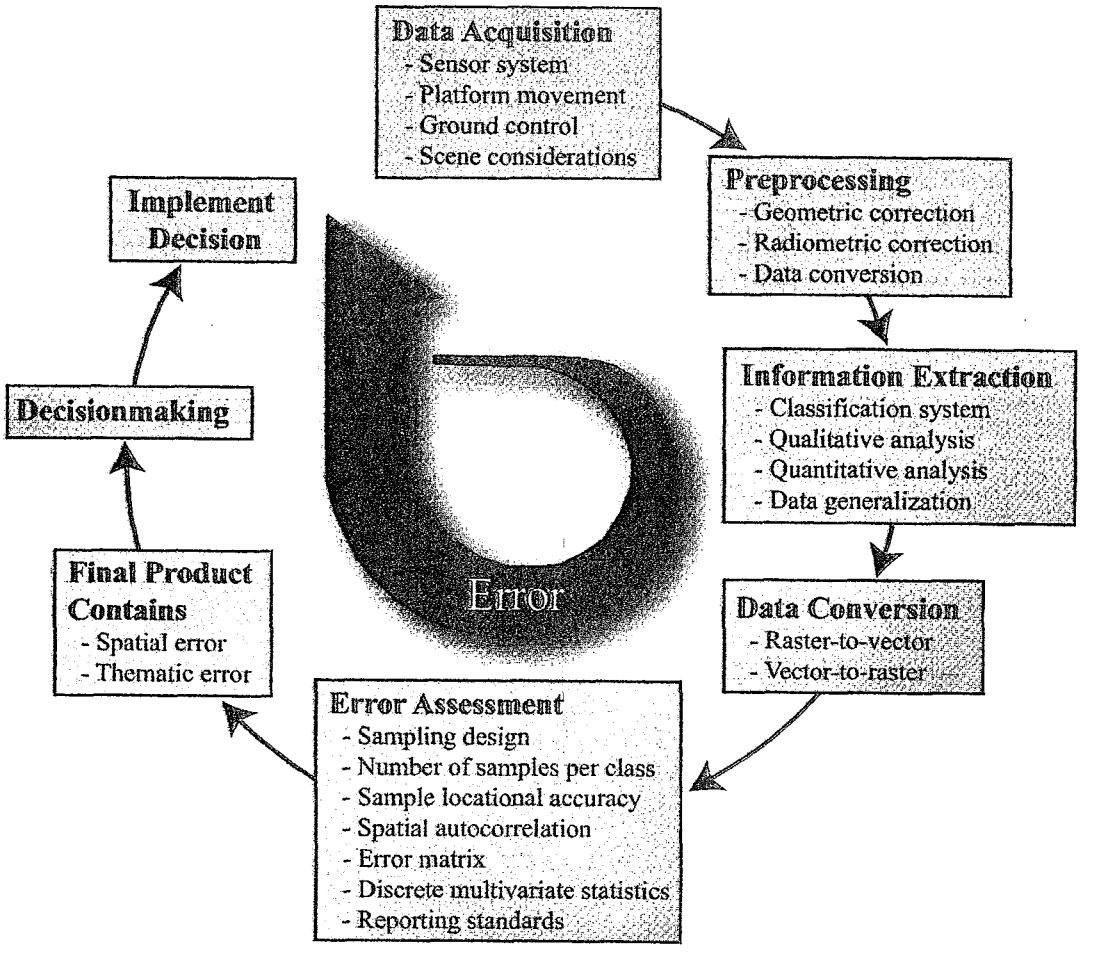


Figure 8-1 Source of Error in Remote Sensing-derived Thematic Products. (From Jensen, 2005).

8.2 General Steps to Assess the Accuracy of Thematic Information Derived from Remotely Sensed Data

Jensen (2005) suggested general steps to assess the accuracy of thematic information derived from remotely sensed data as shown in Figure 8-2 included:

1. State the nature of the thematic accuracy assessment problem
2. Select methods of thematic accuracy assessment
3. Compute total number of observations required in the sample
4. Select sampling design (scheme)
5. Obtain ground reference data at observation locations using a response design
6. Error matrix creation and analysis
7. Accept or reject previously stated hypothesis
8. Distribute results if accuracy is acceptable

Detail of main steps to assess the accuracy of thematic information derived from remotely sensed data will be explained in following section included:

- Error Matrix
- Sample Size
- Sampling Design
- Evaluation of Error Matrix.

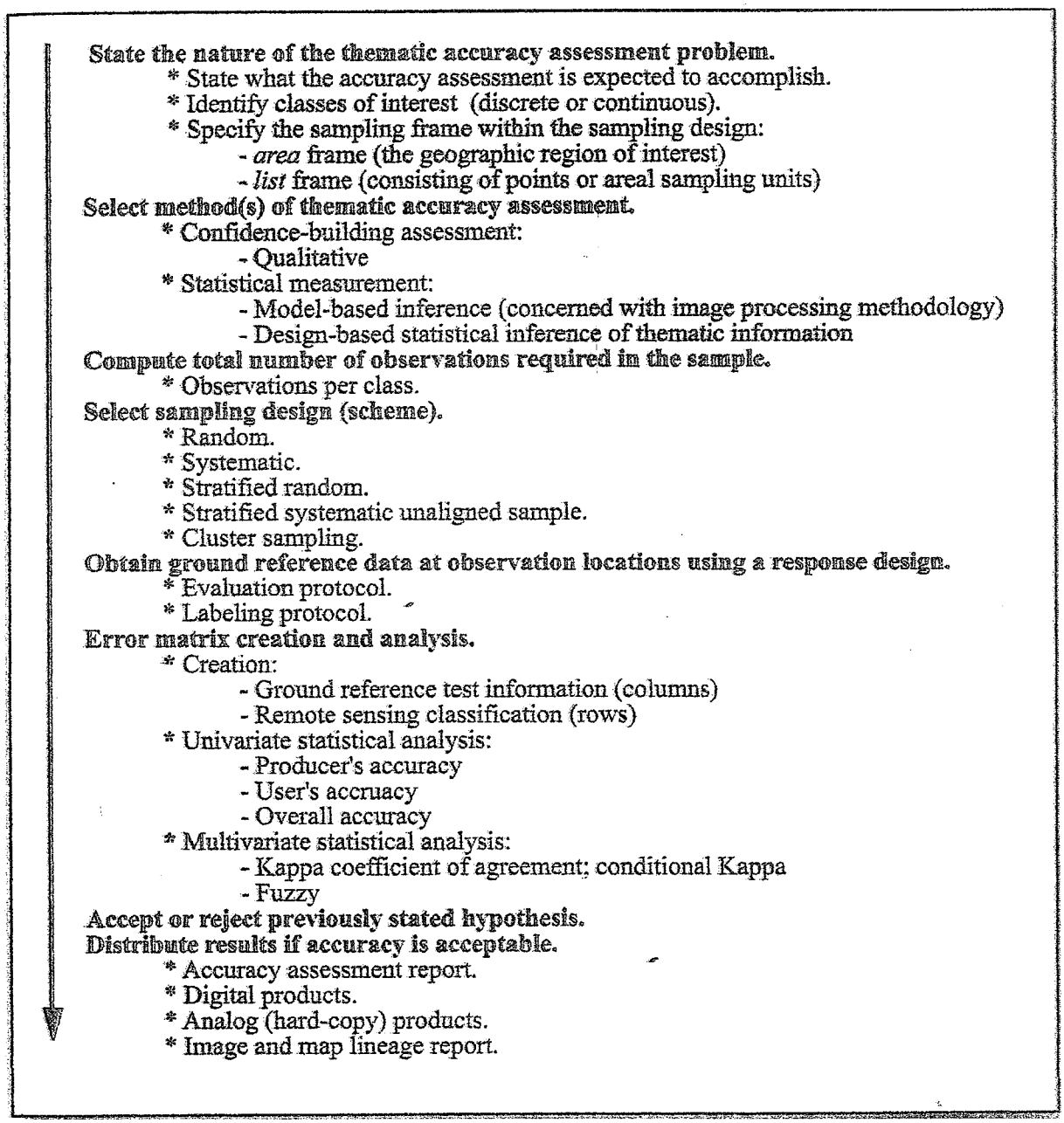


Figure 8-2 General steps to assess the accuracy of thematic information derived from remotely sensed data. (From Jensen, 2005).

8.2.1 The Error Matrix

To correctly perform a classification accuracy (or error) assessment, it is necessary to systematically compare two sources of information:

1. pixels or polygons in a remote sensing-derived classification map and
2. ground reference test information (which may in fact contain error).

The relationship between these two set of information is commonly summarized in an error matrix (contingency table or confuse matrix). Indeed, the error matrix provides the basis on which to both describe classification accuracy and characterize errors, which may help refine the classification or estimates derived from it.

A typical of error matrix is shown in Table 8-1. The columns of the matrix represent the ground reference test information and the rows correspond to the classification generated from analysis of the remotely sensed data. The total number of samples examined is N. The *diagonal of the matrix* summarizes those pixels or polygons that were assigned to the correct class. Every error in the remote sensing classification relative to the ground reference information is summarized in the *off-diagonal cells of the matrix*. Each error is both an omission from the correct category and a commission to a wrong category. The column and row totals around the margin of the matrix are used to compute errors of inclusion (commission errors) and errors of exclusion (omission errors). The outer row and column totals are used to compute producer's and user's accuracy.

Table 8-1: Characteristic of error matrix

		Ground Reference Test Information				
	Class	1	2	3	K	Row total
Remote Sensing Classification	1	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,k}$	x_{1+}
	2	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,k}$	x_{2+}
	3	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$	$x_{3,k}$	x_{3+}
	K	$x_{k,1}$	$x_{k,2}$	$x_{k,3}$	$x_{k,k}$	x_{k+}
	Column total	x_{+1}	x_{+2}	x_{+3}	x_{+k}	N

8.2.2 Sample Size

The actual number of ground reference test samples to be used to assess the accuracy of individual categories in a remote sensing classification map is a very important consideration. Typically, a design-based inferential framework is adopted in accuracy assessment programs. With such a framework, an appropriately sized sample may be estimated using *conventional statistics*. Some analysts use an equation based on the *binomial distribution* or the normal approximation to the binomial distribution to compute the required sample size. Others suggest that a *multinomial distribution* to used to determine the sample size because we are usually investigating the accuracy of multiple classes of information on land-cover map (Jensen, 2005).

8.2.2.1 Sample Size Based on Binomial Probability Theory

Fitzpatrick-Lins (1981) suggested that the sample size N to be used to access the accuracy of a land-use classification map to determined from the formula for the binomial probability theory:

$$N = \frac{Z^2(p)(q)}{E^2} \qquad (8\text{-}1)$$

Where

P is the expected percent accuracy of the entire map,

q is 100 – p

E is the allowable error

Z is 2 from the standard normal deviate of 1.96 for the 95 % two side confidence level.

For a sample for which the expected accuracy is 85% at the allowable error of 5%, the number of the point necessary for reliable results is:

$$N = \frac{2^2(85)(15)}{5^2} = 203$$

8.2.2.2 Sample Size Based on Multinomial Distribution

Analyst usually makes thematic maps that contain multiple classes. Therefore, some scientists prefer to use equation based on a *multinomial distribution* to determine the sample size necessary to assess classification accuracy. Sample size (*N*) derived from a multinomial distribution is based on the equation (Congalton and Green, 1999):

$$N = \frac{B\Pi_i(1-\Pi_i)}{b_i^2} \qquad (8\text{-}2)$$

where:

Π_i is the portion of a population in the *ith* class out of *k* classes that has the proportion closest to 50%

B_i is the desired precision (e.g., 5%) for the class

B is the upper (α/k) X 100 percentile of the chi square ($\chi 2$) distribution with 1 degree of freedom

k is the number of classes.

For example, suppose a land-cover map contains eight classes ($k=8$) and we know that class Π_i occupies approximately 30% of the map area and that this portion is closest to 50%. Furthermore, we desire a level of confidence of 95% and a precision (b_i) of 5%. B is determine from $\chi 2$ table with 1 degree of freedom and $1-\alpha/k$ as $\chi 2_{(1,0.99375)} = 7.568$:

$$1-\frac{\alpha}{k}=1-\frac{0.05}{8}=0.99375$$

$$N=\frac{7.568(0.30)(1-0.30)}{0.05^2}=\frac{1.58928}{0.0025}=636 \quad \text{samples}$$

Therefore, in this example 636 samples should be randomly selected to adequately fill the error matrix. Approximately 80 samples per class are required (e.g., 8 x 80 = 640).

If we have no idea about the proportion of any of the class in the land-cover map, then we can use the worst-case multinomial distribution algorithm where we assume that one class occupied 50% of the study area (Congalton and Green, 1999):

$$N=\frac{B}{4b^2} \qquad (8\text{-}3)$$

Therefore, we desire a level of confidence of 95% and a precision (b_i) of 5% for all k classes yields:

$$N=\frac{7.568}{4(0.05^2)}=757 \quad \text{samples}$$

Thus, 757 random samples would have to be obtained because we did not have prior knowledge about the true proportion of any of the *k* classes in the worse-case scenario.

It is not always possible to obtain such large numbers of random samples. A balance between what is statistically sound and what is practicably attainable must be found. Congalton and Green (1999) suggested that a good rule of thumb is to collect a minimum of 50 samples for each land-cover class in the error matrix. If the area is especially large (e.g., > 1 million ha) or the

classification has large number of land-use categories (e.g., > 10 classes), the minimum number of sample should be increase to 75 or 100 samples per class.

8.2.3 Sampling Design

There are basically five sampling designs used to collect ground reference test data for accuracy assessment (Jensen, 2005) include (1) Random sampling, (2) Systematic sampling, (3) Stratified random sampling, (4) Stratified systematic unaligned sampling, and (5) Clustering sampling.

8.2.3.1 Simple Random Sampling

Simple *random sampling* without replacement always provides adequate estimates of the population parameters, provides the sample size is sufficient (Figure 8-3a) (Congalton, 1988). Normally a random number generator is used to identify random x, y coordinate within the study area. This method ensured that the members of a population are drawn independently with equal probabilities of being selected and that the selection of one sample does not influence the selection of any other sample. Unfortunately, simple random sampling may undersample small but possibly very important classes unless the sample size is significantly large. The random samples may also be located in inhospitable or cross-denied location. Finally, some clustering of the random sample points often results and, as a consequence, important spatial properties of the population may be overlooked.

8.2.3.2 Systematic Sampling

To avoid the shortcoming of random sampling regarding the coverage of an area, a *systematic sampling* plan may be used. This procedure requires that it be possible to work through the study area in some consistent and orderly manner such as along a line of coordinate points. The sampling is normally initiated by the random selection of the starting point x, y coordinates (A) as shown in Figure 8-3b. Unfortunately, if there are periodicities in the data being collected (e.g., the ridge and valley topography mountain), the regularly spaced points could hit the same on a cycle time and again, yielding biased sample estimates of the population. Systematic sampling should be used with caution because it may overestimate the population parameters (Congalton, 1988).

8.2.3.3 Stratified Random Sampling

Many remote sensing analysts prefer stratified random sampling by which a minimum number of samples are selected from each strata after the thematic map has been prepared. Stratified random sampling involves two steps. First, the study is created by extracting only pixels associated with a specific class. Some analysts prefer to make a mask out all unwanted classes to create the stratified. Sample locations are then randomly distributed throughout the geographic strata as shown in Figure 8-3c. The advantage of stratified random sampling is that all strata, no matter how small their proportion of the entire study area, will have samples allocated to them for error evaluation process. The drawback with stratified random sampling is that one has to wait until the thematic map has been completed in order to allocate samples to the various land-cover strata.

8.2.3.4 Stratified Systematic Unaligned Sampling

Stratified systematic unaligned sampling combines randomness and stratification with a systematic interval. It introduces more randomness than just beginning with a random x,y coordinate for the first sample in each stratum. Figure 8-3d demonstrates how a stratified systematic unaligned sample is created. First, point A is selected at random. The x-coordinate of A is then used with a new random y-coordinate to located B, a second random y-coordinate to locate E, and so on across the top row of the stratum. By a similar process the y-coordinate of A is used in combination with random x-coordinate to locate point C and all successive point in the first column of the stratum. The random x-coordinate of C and y-coordinate of B are then used to locate D, of E and F to locate G, and so on until all strata have sample elements (King, 1969).

8.2.3.5 Clustering Sampling

Sometimes it is difficult to go into the field to randomly selected locations and obtain the required information. Therefore, some have suggested that several samples can be collected at a single random location. Unfortunately, each pixel in the cluster is not independent of the others. Therefore, Congalton (1988) warns that clusters no larger than 10 pixels and certainly not larger than 25 pixels should be used because of the lack of information added by each pixel beyond these cluster size (Figure 8-3e).

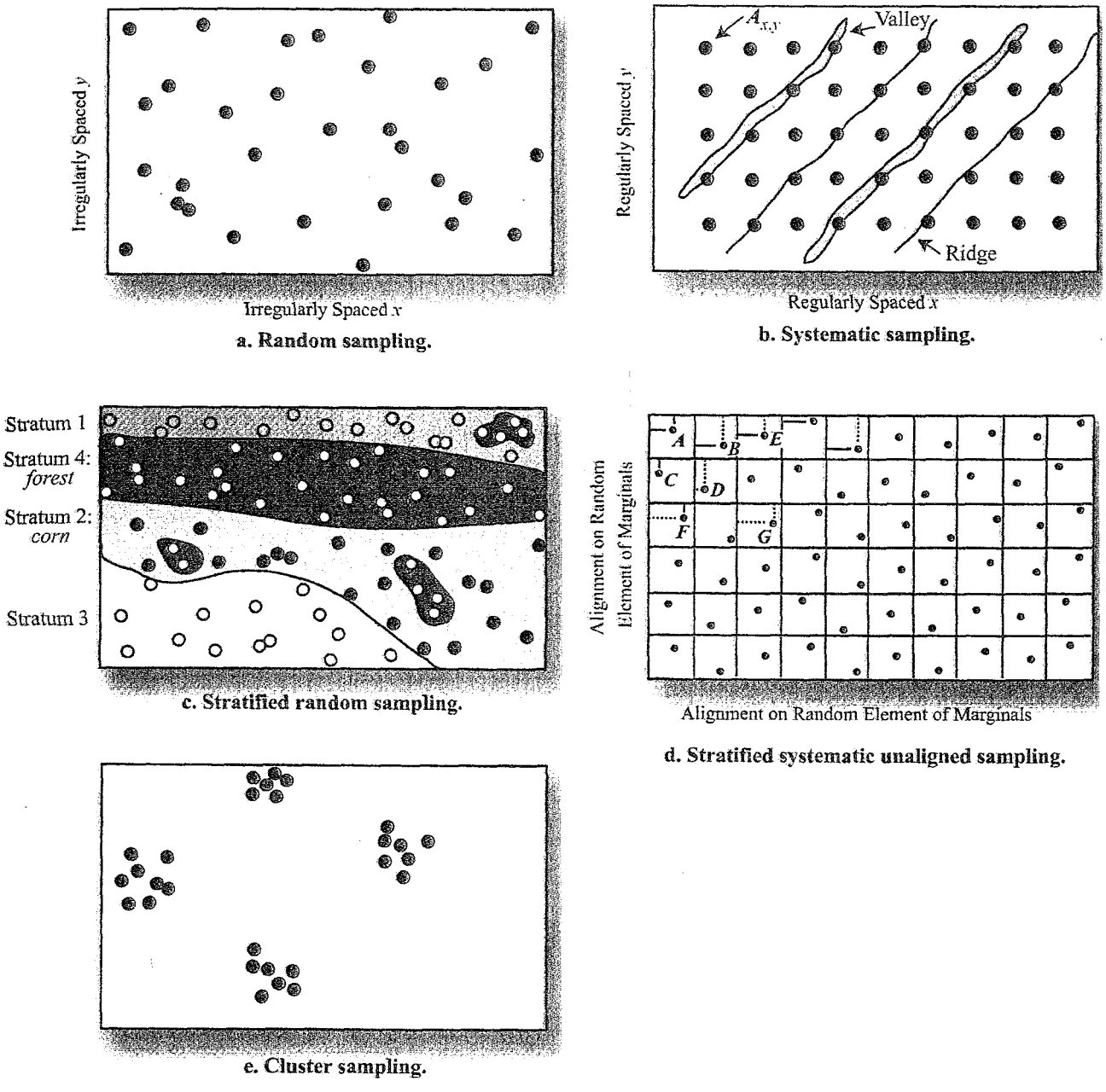


Figure 8-3 Geographic Sampling methods (From Jensen, 2005).

8.2.4 Evaluation of Error Matrices

After the ground reference test information has been collected from the randomly located sites, the test information is compared pixel by pixel (or polygon by polygon when the remote sensor data are visually interpreted) with the information in the remote sensing-derived classification map. Agreement and disagreement are summarized in the cell of the error matrix. Then, information in the error matrix may be evaluated using:

- simple descriptive statistics, or
- multivariate analytical statistical techniques: Kappa Analysis.

8.2.4.1 Descriptive Evaluation of Error Matrices

The *overall accuracy* of the classification map is determined by dividing the total correct pixels (sum of the major diagonal) by the total number of pixels in the error matrix (*N*). Computing the accuracy of individual categories, however, is more complex because the analyst has the choice of dividing the number of pixels in the corresponding row and column. Traditionally, the total number of correct pixels in a category as divided by the total number of the pixels of that category as derived from the reference data (i.e., the column total). This statistics indicates the probability of a reference pixel being correctly classified and is a measure error of exclusion (*omission error*). This statistics is called the *producer's accuracy* because the producer (the analyst) of the classification is interested in how well a certain area can be classified.

If the total number of correct pixels in a category is divided by the total number of pixels that were actually classified in that category, the result is a measure of inclusion (*commission error*). This measures, called the *user's accuracy* or reliability, is the probability that a pixel classified on the map actually represents that category on the ground (Story and Congalton, 1986). The example of error matrix and their accuracy computation is shown in Table 8-2.

Table 8-2: Error matrix and their accuracy assessment by simple descriptive statistics.

		Reference Data					
		Residential	**Commercial**	**Wetland**	**Forest**	**Water**	**Row Total**
Remote Sensing Classification Data	**Residential**	70	5	0	13	0	88
	Commercial	3	55	0	0	0	58
	Wetland	0	0	99	0	0	99
	Forest	0	0	4	37	0	41
	Water	0	0	0	0	121	121
	Column Total	73	60	103	50	121	407

Calculation:

Overall Accuracy = 382/407 = 93.86%

Producer's Accuracy / Omission Error	**User's Accuracy / Commission Error**
Residential = 70/73 = 96%/ 4% Omission error	**Residential = 70/88 = 80%/ 20% Commission error**
Commercial = 55/60 = 92%/ 8% Omission error	**Commercial = 55/58 = 95%/ 5% Commission error**
Wetland = 99/103 = 96%/ 4% Omission error	**Wetland = 99/99 = 100%/ 20% Commission error**
Forest = 37/50 = 74%/ 26% Omission error	**Forest = 37/41 = 90%/ 10% Commission error**
Water = 121/121 = 100%/ 0% Omission error	**Water = 121/121 = 100%/ 0% Commission error**

8.2.4.2 Discrete Multivariate Analytical Techniques: Kappa Analysis

Discrete multivariate techniques have been used to statistically evaluate the accuracy of remote sensing-derived classification maps and error matrices since 1983 The techniques are appropriate because remotely sensed data are discrete rather than continuous and are also binomially or multinomially distributed rather than normally distributed. However, it is important to remember that there is not a single universally accepted measure of accuracy but instead a variety of indices, each sensitive to different features (Jensen, 2005).

The Kappa analysis is a discrete multivariate technique used in accuracy assessment for statistically determining if on error matrix is significantly different than another. The result of performing a Kappa analysis is KHAT statistic (an estimate of Kappa), which is another of agreement of accuracy. This measure of agreement is based on the difference between the actual agreement in the error matrix (i.e., the agreement between the remotely sensed classification and the reference data as indicated by the major diagonal) and the change agreement which indicated by the row and column totals (marginals) (Congalton and Green, 1999). KHAT is computed:

$$\hat{K} = \frac{N\sum_{i=1}^{k} x_{ii} - \sum_{i=1}^{k}(x_{i+} \times x_{+1})}{N^2 - \sum_{i=1}^{k}(x_{i+} \times x_{+1})} \qquad (8\text{-}4)$$

where

- k is the number of rows (e.g. land-cover classes) in the matrix,
- x_{ii} is the number of the observation in row i and column i
- x_{i+} is the marginal totals for row i
- x_{+i} is the marginal totals for column i
- N is the total number of observations. (See example in Table 8-3.

Table 8-3: Error matrix and their accuracy assessment by Kappa Analysis

		Reference Data					
		Residential	**Commercial**	**Wetland**	**Forest**	**Water**	**Row Total**
Remote Sensing Classification Data	**Residential**	**70**	**5**	**0**	**13**	**0**	**88**
	Commercial	**3**	**55**	**0**	**0**	**0**	**58**
	Wetland	**0**	**0**	**99**	**0**	**0**	**99**
	Forest	**0**	**0**	**4**	**37**	**0**	**41**
	Water	**0**	**0**	**0**	**0**	**121**	**121**
	Column Total	**73**	**60**	**103**	**50**	**121**	**407**

$$\hat{K} = \frac{N\sum_{i=1}^{k} x_{ii} - \sum_{i=1}^{k}\left(x_{i+} \times x_{+1}\right)}{N^2 - \sum_{i=1}^{k}\left(x_{i+} \times x_{+1}\right)}$$

Where N = 407

$\sum_{i=1}^{k} x_{ii}$ = (70 + 55 + 99 + 37 +121) = 382

$\sum_{i=1}^{k}\left(x_{i+} \times x_{+1}\right)$ = (88 x 73) +(58 x 60) + (99 x 103) + (41 x 50) + (121 x 121) = 36,792

Thus, $\hat{K} = \frac{(407 \times 382) - 36792}{407^2 - 36792} = \frac{155474 - 36792}{165649 - 36792} = \frac{118682}{128857} = 92.1\%$

Jensen (2005) stated that Landis and Koch (1977) concluded the values of KHAT as follows:

- KHAT values > 0.80 (i.e., >80 %) represent strong agreement or accuracy between the classification map and the ground reference information.
- KHAT values between 0.40 and 0.80 (i.e., 40% to 80 %) represent moderate agreement or accuracy between the classification map and the ground reference information.
- KHAT values < 0.40 (i.e., <40 %) represent poor agreement or accuracy between the classification map and the ground reference information.

Congalton (1991) suggests that the overall accuracy, normalized accuracy and KHAT be computed for each matrix to "glean as much information from the error matrix as possible." Computation of the KHAT statistic can be used (1) to determine whether the results presented in the error matrix are significantly better a random result (i.e., a null hypothesis of KAHT = 0) or (2) to compare two similar matrices consisting of identical categories to determine if they are significant different.

In addition to computing the Kappa coefficient for an entire error matrix, Congalton and Green (1999) suggest that it may be useful to look at the agreement for an individual category within the matrix. Individual category agreement can be tested using the conditional Kappa coefficient agreement for the i^{th} category is given by:

$$\hat{K}_i = \frac{N(x_{ii}) - (x_{i+} \times x_{+i})}{N(x_{i+}) - (x_{i+} \times x_{+i})} \quad (8\text{-}5)$$

where

x_{ii} is the number of observations correctly classified for a particular category (summarized in the diagonal of the matrix),

x_{i+} is marginal totals for row i associated with category,

x_{+i} is marginal totals for column i associated with category,

N is the total number of observations in the entire matrix.

For example, the conditional Ki coefficient of agreement for the residential land-use class dataset in Table 8-3 is:

$$\hat{K}_{Residential} = \frac{407(70) - (88 \times 73)}{407(88) - (88 \times 73)} = \frac{28490 - 6424}{35816 - 6424} = 0.75$$

8.2.4.3 Fuzzification of the Error Matrix

Congalton and Green (1999) introduced additional methods to incorporate fuzzy logic into an error assessment using a fuzzifield error matrix. The first method involves as accepting as correct +- class along the diagonal in the error matrix. This assumes that the data represent a continuum. The marginal column and row totals take this enlargement of the diagonal (i.e., fuzzification) into account, as does overall classification accuracy.

Table 8-4 present the results of a set of fuzzy rule applied for forest crown closure classification. In this case, the rules were defined as follows:

- Class 1 was defined as always 05 crown closure. If the reference data indicated a value of 0%, then only an image classification of 0% was accepted.
- Class 2 was defined as acceptable if the reference data was 5 % of that of the image classification. In other words, if the reference data indicates that a sample has 15% crown closure and the image classification put it Class 2, the answer would not be absolutely correct, bur acceptable.

- Classes 3 through 6 were as acceptable if the reference data were within 10% of that of the image classification. In other words, a sample classified as Class 4 on the image but found to be 55% crown closure on the reference data would be considered acceptable.

As a result of these rules, off-diagonal elements in the matrix contain two separate values. The first value represents those that, although not absolutely correct, are acceptable. Therefore, in order to compute the overall, producer's and user's accuracies, the values along the major diagonal and those deemed acceptable (i.e., those in the first value) in the off-diagonal elements are combined. This combination of absolutely correct and acceptable answers results in an overall accuracy of 64%. This overall accuracy is significantly higher than the original error matrix with an overall accuracy of 40% as shown in Table 8-5.

Table 8-4: Error matrix and their accuracy assessment with Fuzzy Logic Rule.

Ground Reference

Classification Data	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Row Total
Class 1	2	6,3	1	2	1	1	15
Class 2	0,2	8	2,1	6	1	1	21
Class 3	0	2,1	3	4,0	9	1	20
Class 4	0	0	0,2	8	5,2	10	27
Class 5	0	1	2	1,0	6	12,4	26
Class 6	0	0	0	0	2,1	31	34
Column Total	4	21	11	21	27	60	144

Overall accuracy = 92/144 = 64%

Producer's Accuracy

Class 1 = 2/4 = 50%
Class 2 = 16/21 = 76%
Class 3 = 5/11 = 45%
Class 4 = 13/21 = 62%
Class 5 = 13/27 = 48%
Class 6 = 43/60 = 72%

User's Accuracy

Class 1 = 8/16 = 50%
Class 2 = 10/21 = 48%
Class 3 = 9/20 = 45%
Class 4 = 13/27 = 48%
Class 5 = 19/26 = 73%
Class 6 = 33/34 = 97%

From Congalton and Green, 1999.

Table 8-5: Error matrix and their accuracy assessment without Fuzzy Logic Rule.

Ground Reference

Classification Data	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Row Total
Class 1	2	9	1	2	1	1	15
Class 2	0	8	3	6	1	1	21
Class 3	0	3	3	4	9	1	20
Class 4	0	0	2	8	7	10	27
Class 5	0	1	2	1	6	6	26
Class 6	0	0	0	0	3	31	34
Column Total	4	21	11	21	27	60	144

Overall accuracy = 58/144 = 40%

Producer's Accuracy

Class 1 = 2/4 = 50%

Class 2 = 8/21 = 38%

Class 3 = 3/11 = 27%

Class 4 = 8/21 = 38%

Class 5 = 6/27 = 22%

Class 6 = 31/60 = 52%

User's Accuracy

Class 1 = 2/16 = 13%

Class 2 = 8/21 = 38%

Class 3 = 3/20 = 15%

Class 4 = 8/27 = 30%

Class 5 = 6/26 = 23%

Class 6 = 31/34 = 91%

From Congalton and Green, 1999

Chapter 9: Digital Change Detection

Biophysical materials and human-made features on the surface of Earth are inventories using remote sensing and in situ techniques. Some of the data are fairly static; they do not change over time. Conversely, some biophysical materials and human-made features are dynamic, changing rapidly. It is important that such changes be inventories accurately so that the physical and human processes at work can be more fully understood. In fact, it is believed that land-use and land-cover change is a major component of global change with an impact perhaps greater than that of climate change (Skole, 1994; Foody, 2001).

This chapter will review how change information is extracted from digital remotely sensed data. It summarized the remote sensor system and environment parameters that must be considered when change detection takes place. Several of the most widely used change detection algorithms are introduced and demonstrated.

9.1 Step Required to Perform Change Detection

Lunetta (1999) suggested general data processing elements for a remote sensing change detection application project or scientific study as shown in Figure 9-1. The analytical methods employed to accomplish each individual element will be vary considerably based on the specific goals and objectives defined using the project formulation process (Figure 9-2). The final result is the development of a data processing flow diagram that sequentially details all analysis steps and data processing time line (optional). The data processing flow diagram is particular important to facilitate the implementation of complex data analysis projects, especially those involving numerous investigations or technicians.

Jensen (2005) suggested the general steps required to perform change detection from digital multispectral remote sensor data were described in Figure 9-3 including:

1. State the nature of the change detection problem
2. Consideration of significance when performing change detection
3. Process remote sensor data to extract change information
4. Perform accuracy assessment
5. Accept or reject previously stated hypothesis
6. Distribute results if accuracy is acceptable

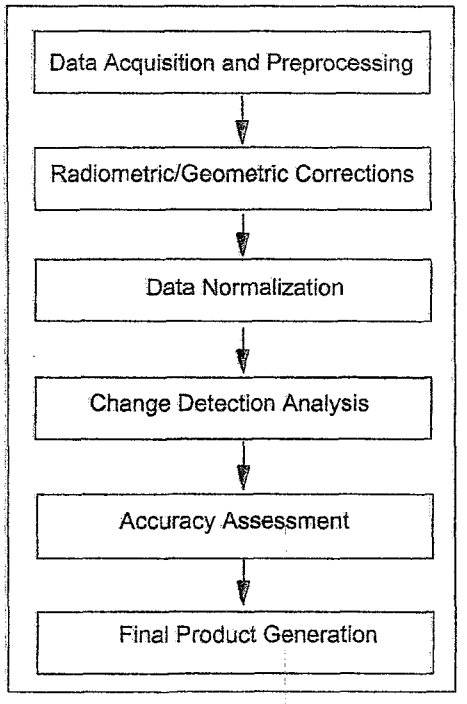


Figure 9-1: General data processing element for a remote sensing change detection application project or scientific study (From Lunetta, 1999)

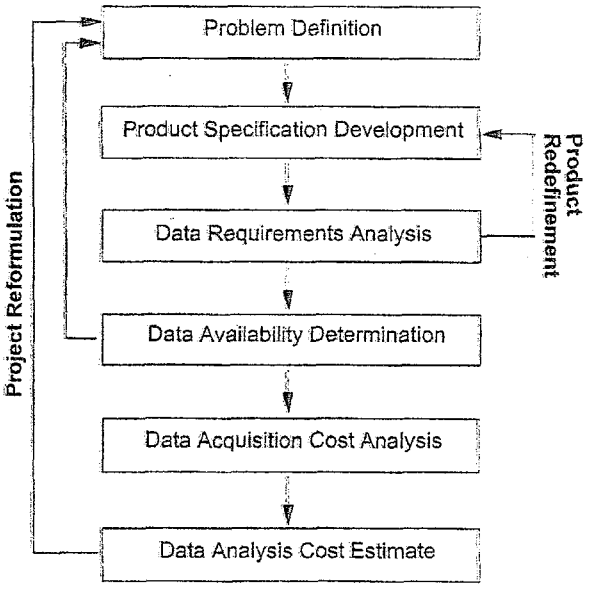


Figure 9-2: Sequential steps outlining the project formulation process for remote sensing change detection project (From Lunetta 1999)

Detail of important steps to conduct digital change detection of remotely sensed data will be here explained based on Jensen (2005) in following section including:

- Remote sensing considerations;
- Environmental characteristics considerations;
- Selection of a change detection algorithm.

State the nature of the change detection problem.
* Specify change detection geographic region of interest.
* Specify change detection time period (e.g., daily, seasonal, yearly).
* Define the classes of interest in a classification system.
* Select hard and/or fuzzy change detection logic.
* Select per-pixel or object-oriented change detection.

Considerations of significance when performing change detection.
* Remote sensing system considerations:
  - Spatial, spectral, temporal, and radiometric resolution
* Environmental considerations:
  - Atmospheric conditions
  - Soil moisture conditions
  - Phenological cycle characteristics
  - Tidal stage, etc.

Process remote sensor data to extract change information.
* Acquire appropriate change detection data:
  - *In situ* and collateral data
  - Remotely sensed data:
    - Base year (time n)
    - Subsequent year(s) (time $n - 1$ or $n + 1$)
* Preprocess the multiple-date remote sensor data:
  - Geometric correction
  - Radiometric correction (or normalization)
* Select change detection algorithm.
* Apply appropriate image classification logic if necessary:
  - Supervised, unsupervised, hybrid
* Perform change detection using GIS algorithms:
  - Highlight selected classes using change detection matrix
  - Generate change-map products
  - Compute change statistics

Perform accuracy assessment.
* Select method:
  - Qualitative confidence building
  - Statistical measurement
* Determine number of samples required by class.
* Select sampling scheme.
* Obtain ground reference test information.
* Create and analyze change detection error matrix:
  - Univariate and multivariate statistical analysis

Accept or reject previously stated hypothesis.
Distribute results if accuracy is acceptable.

Figure 9-3: The general steps used to perform digital change detection of remotely sensed data (From Jensen, 2005).

9.2 Remote sensor system considerations

Failure to understand the impact of the various parameters on the change detection process can lead to inaccurate results. Ideally, the remotely sensed data used to perform change detection is acquired by a remote sensor system that holds the following resolutions constant: temporal, spatial (and looking angle), spectral, and radiometric.

A. Temporal Resolution

Two important temporal resolutions should be held constant during change detection using multiple dates of remotely sensed data. First, the data should be obtained from a sensor system that acquired data at approximately the same time of day. This eliminate diurnal Sun angle effects that can cause anomalous difference in the reflectance properties of the remotely sensed data. Second, whenever possible it is desirable to use remotely sensed data acquired on anniversary dates. Using anniversary date imagery minimizes the influence of season Sun-angle and plant phenological differences that can negatively impact a change detection project (Jensen et al., 1993)

B. Spatial Resolution and Look Angle

Accurate spatial registration of at least two images is essential for digital change detection. Ideally, the remotely sensed data are acquired by a sensor system that collects data with the same instantaneous field of view on each date. It is possible to perform change detection using data collected from two different sensor systems with different IFOVs, for example, Landsat TM data (30x30 m) for date 1 and SPOT HRV XS data (20x20 m) for date 2. In such case, it is usually necessary to decide on a representative minimum mapping unit (e.g., 20x20 m) and then resample both datasets to this uniform pixel size. Geometric rectification algorithms are used to register the images to a standard map projection. Rectification should result in the two images having a root mean square error (RMSE) of < 0.5 pixel. Misregistration of the two images may result in the identification of spurious areas of change between datasets.

C. Spectral Resolution

A fundamental assumption of digital change detection is that a difference exists in the spectral response of a pixel on two dates if the biophysical materials within the IFOV have changed between dates. However, different sensor systems do not record energy in exactly the same portions of the electromagnetic spectrum (i.e., bandwidth). Ideally, the same sensor system is used to acquire imagery on multiple dates. When this is used to acquire imagery on multiple dates. When this not possible, the analyst should select bands that approximate one another. For example, Landsat MSS band 4 (green), 5 (red), and 7 (NIR), and SPOT bands 1 (green), 2 (red), and 3 (NIR) can be used successfully with Landsat ETM+ bands 2 (green), 3 (red), and 4 (NIR).

D. Radiometric Resolution

Ideally, the sensor systems collect the data at the same radiometric precision on both dates. When radiometric resolution of data acquired by one system (e.g., Landsat MSS with 6-bit data) is compared with data acquired by a higher radiometric resolution instrument (Landsat TM with 8-bit data), the lower-resolution data (e.g., 6 bits) should be decompressed to 8 bit for change detection purpose. The precision of decompressed brightness values can never be better than the original, non-compressed data. Ideally, the brightness values associated with both dates of imagery are converted to apparent surface reflectance, which eliminates the problem.

9.3 Environmental Characteristics Considerations

Failure to understand the impact of various environmental characteristics on the remote sensing change detection process can lead to inaccurate results. When performing change detection, it is desirable to hold environmental variables as constant as possible. It should considered:

- Atmospheric Conditions
- Soil Moisture Conditions
- Phenological Cycle Characteristics
- Vegetation Phenology
- Urban-Suburban Phenological Cycles
- Effects of Tidal Stages on Change Detection

A. Atmospheric Conditions

There should be no clouds (including stratus) or extreme humidity on the days that remote sensing data are collected. Even a thin haze can alter spectral signatures in satellite images enough to create the false impression of spectral change between two date. Areas obscured by clouds or affected by cloud shadow will filter through the entire change detection product.

Assuming no cloud cover, the use of anniversary dates helps to ensure general, seasonal agreement between the atmospheric conditions on the two dates. If dramatic difference exist in the atmospheric conditions on the n dates of imagery to be used in the change detection process, it may be necessary to remove the atmospheric attenuation in the imagery. For mountainous areas, topographic effects may also have to be removed (Civco, 1989). If it is not possible to perform an absolute radiometric corrected, then image-to-image normalization might be a viable alternative.

B. Soil Moisture Conditions

Ideally, the soil moisture conditions should be identical for the n dates of imagery use in a change detection project. Extremely wet or dry conditions on one of the date can cause serious

change detection problems. When soil moisture differences between dates are significant for only certain parts of the study area, it may be necessary to stratify (cut out) those affected areas and perform a separate change detection analysis, which can be added back in the final stages of the project.

C. Phenological Cycle Characteristics

Natural ecosystems go through repeatable, predictable cycles of development. Human being often modify the landscape in repeatable, predictable stages. These cycles of predictable development are often referred to as phenomenological or phenological cycles. Image analysts use these cycles to identify when remotely sensed data should be collected to obtain the maximum amount of usable change information. Therefore, analysts must be intimately familiar with biophysical characteristics of the vegetation, soils, and water constituents of ecosystems and their phenological cycles.

D. Vegetation Phenology

Vegetation grows according to relatively predictable diurnal, seasonal, and annual phenological cycles. Obtaining near-anniversary image greatly minimizes the effects of seasonal phenological differences that may cause spurious change to be detected in the imagery. Ideally, monoculture crops are planted at approximately the same time of year on the two dates of imagery. A month lag in planting date between fields of the same crop can cause serious change detection error. Second, the monoculture should be the same species. In addition, changes in row spacing and direction can have an impact. These observations suggest that the analyst must know the crop's biophysical characteristics as well as the cultural land-tenure practices in the study area so that the most appropriate remotely sensed data can be selected for change detection. Natural vegetation ecosystem such as wetland aquatic plants, forests, and rangeland have unique phenological cycles as shown in Figure 9-4.

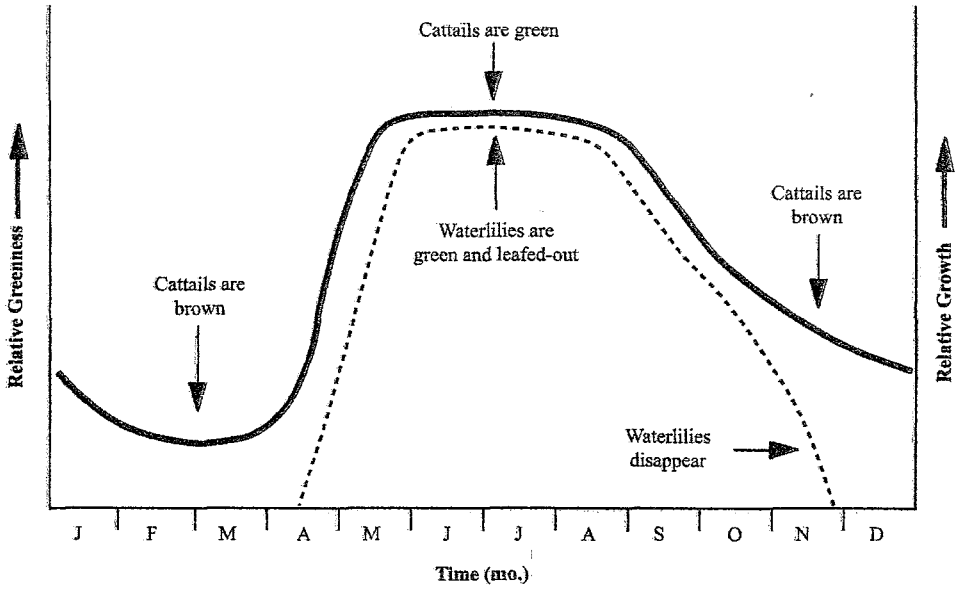


Figure 9-4: Phenological cycle of Cattails and Waterlilies in Par Pond in SC. (From Jensen, 2005).

E. Urban-suburban Phenological Cycles

Urban-suburban landscape also have phenological cycles. Most novice image analysts assume that change detection in the urban-rural fringe will capture the residential developed in the two most important stages: rural undeveloped land and completely developed residential. Jensen (1981) identified 10 stages of residential development taking place in this region based on evidence of clearing, subdivision, transportation, buildings, and landscaping. Many of these stages may appear spectrally similar to other phenomena. Therefore, the analyst must be intimately aware of the phenological cycle of all urban phenomena being investigated as well as the natural ecosystems.

In addition, Yuan et al (1999) introduced some of the complicating factors which tend to make change detection more difficult for scene selection and preprocessing steps which can be used to address these complications for ease in land cover change analyses include:

- Spatial resolution and spectral bandpass differences between images acquired with two sensors complicates the direct comparison of the digital analysis of the data to detect changes (Landsat MSS-four spectral bands with 80 meter pixels and Landsat-TM-seven bands with 30 meter pixels except Band 6). With the spectral bandpass difference, there may not be a direct compatibility of the detectable land cover classes between the two dates of imagery. These complications can be minimized by using data from a single sensor series.

- If clouds are present on one or both dates of imagery it is impossible to detect land cover change between the two dates of imagery. To minimize the effects of clouds, low cloud cover should be one of the primary factors directing the selection of scene for a project. Alternatively, it is possible to composite multiple cloud-impacted scenes to assemble cloud-reduced image products. The compositing to reduce cloud cover can be quite effective if the composited scenes are closely method radiometrically and phonologically.
- Phenological variations in vegetation result in large changes in the reflectance patterns of the land surface. If images of leaf-on and leaf-off conditions are compared digitally, large areas can appear to have been "deforested." This problem can be minimized by selecting scenes from the same time of year for each of the three time periods. Spectral differences in vegetation between wet years and dry years can be quite pronounced, even if images dates are closely matched. In this event it may be necessary to select imagery from alternate years, which are more closely matched in their precipitation pattern. For areas with high cloud cover and sparse satellite data archives, it may not be possible to select scenes that are temporally matched. In this case, the post-classification route to land cover change analysis is indicated.
- Differences in the radiometric performance between sensors or changes in the performance of an individual sensor over time can complicate change detection by inducing spectral differences that are unrelated to land cover changes. This style of effect is best addressed using data from an onboard calibration system to provide gains and offsets for conversion of digital data to radiance units (watt/cm^2/sr/μm). Unfortunately, not every sensor is equipped with onboard calibration systems (i.e., Landsat 1-5). The simplest way to address sensor-related radiometric differences is through scene-to-scene radiometric normalization.
- Variation in solar irradiance, solar zenith angle, and solar azimuth will affect scene brightness levels and the location of shadows, all of which may impact a land cover changes analysis. Those effect can be minimized by selecting scenes from different years at or near the same Julian date in order to match the solar conditions as far as possible. In general, high sun angle (low amount of shadowing) are better than low sun angles for the detection of land cover change.
- Scene-to-scene variations in atmospheric effects (scattering and absorption) can affect scene characteristics sufficiently that they must be considered in evaluating change detection methods. By assuming that the atmospheric effects on the selected scenes are uniform across the entire scene area, it is possible to compensate for scene-to-scene variations in atmospheric effects through atmospheric correction or radiometric

normalization. In some cases even the best "low-cloud" imagery has obvious atmospheric haze or cirrus cloud cover which are not evenly distributed across the scene, confounding simple methods to prepare the data for digital change detection.

- Spatial misregistration of image will tend to reduce the accuracy of any digital change detection effort. Generally, coregistration accuracies should be on the order of half a pixel or less, creating data set which can be analyzed for change with minimal errors due to misregistration.

9.4 Selection of a Change Detection Algorithm

The selection of an appropriate change detection algorithm is very important:

- First, it will have a direct impact on the type of image classification to be performed (if any).
- Second, it will dictate whether important "from-to" change information can be extracted from the imagery.

Jensen (2005) categorized change detection algorithms that commonly used include:

- Write Function Memory Insertion
- Multi-date Composite Image
- Image Algebra
- Post-classification Comparison
- Binary Change Mask Applied to Date 2
- Ancillary Data Source as Date 1
- Spectral Change Vector Analysis
- Chi-square Transformation
- Cross-correlation
- Knowledge-based Vision Systems.

Explanation of each algorithm was summarized based on Jensen (2005) as following:

9.4.1 Change Detection Using Write Function Memory Insertion

A simple powerful method of visual change detection involves the use of the three write function memory (WFM) banks found on the graphics card of every digital image processing system. Basically, individual bands (or derivative products) from multiple dates may be inserted into each of the three WFM banks (red, green, or blue) to identify change in the imagery. This algorithm, advantage and disadvantage are summarized as shown in Figure 9-5, while Figure 9-6 shows an example of change detection using write function memory insertion.

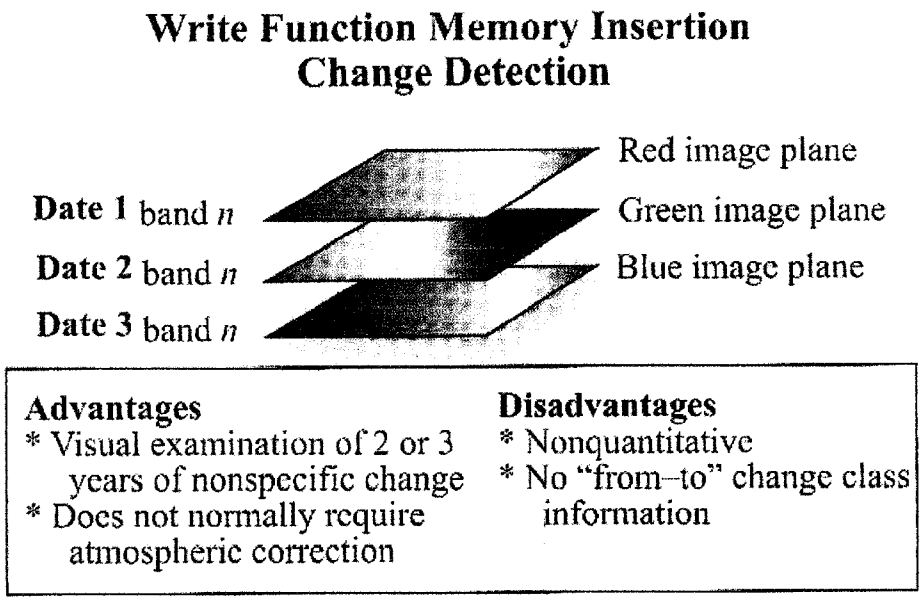


Figure 9-5: Algorithm, advantage and disadvantage of write function memory insertion change detection (Jensen, 2005).

(a) Landsat TM of Lamtakhong Dam, Nakhon Ratchasima, Date 9 February 1999, Band 453 (RGB)

(b) Landsat TM of Lamtakhong Dam, Nakhon Ratchasima, Date 26 March 2004, Band 453 (RGB)

(c) Write function memory insertion change where RGB = Landsat TM Band 5 Date xxx 1999, Landsat TM Band 4 Date 26 March 2004 and Band 5 Date 26 March 2004

Figure 9-6: An example of change detection using write function memory insertion, Lam Takhong Dam, Nakhon Ratchasima, Thailand.

9.4.2 Multidate Composite Image Change Detection

Numerous researchers have rectified multiple dates of remotely sensed imagery (e.g., selected bands of two IKONOS scenes of the same region) and placed them in a single dataset (Figure 9-7). This composite dataset can then be analyzed in a number of ways to extract change information. First, a tradition classification using all n bands may be performed. Unsupervised classification techniques will result in the creation of change and no-change clusters. The analyst must then label the clusters accordingly. Other researchers have subjected the registered composite image dataset to principal component analysis (PCA) to detect change The major components of the derived PCA dataset tend to account for variation in the image data that is not due to land-cover change, and they are termed static components. Minor components tend to enhance spectral contrasts between the two dates, and they are termed change components. as shown in PCA2 in Figure 9-8.

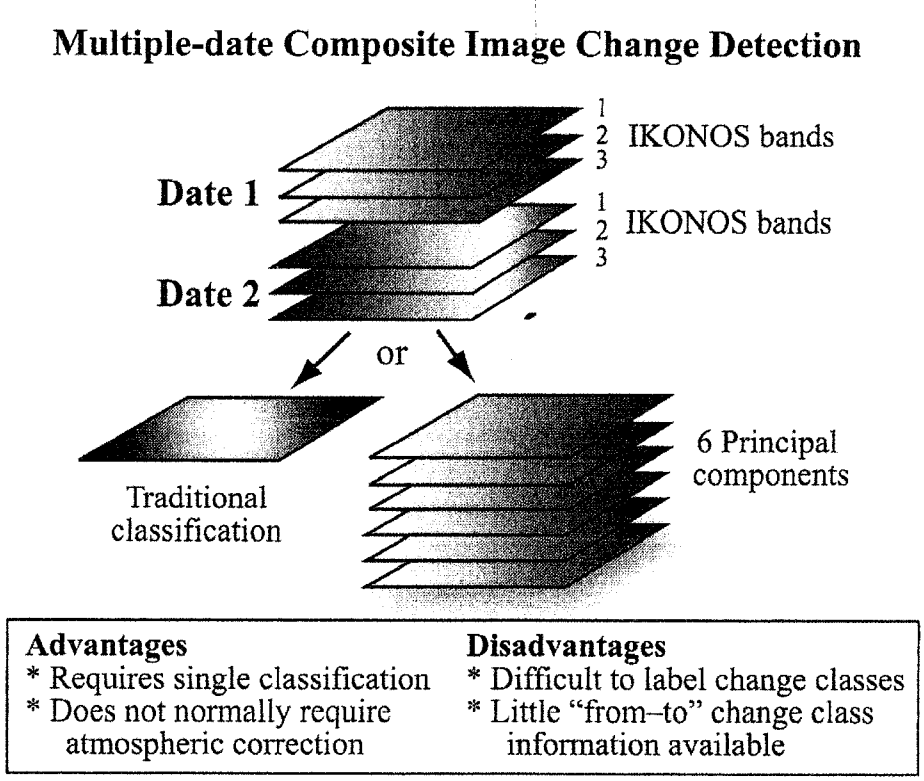


Figure 9-7: Algorithm, advantage and disadvantage of multiple-date composite image change detection (From Jensen, 2005).

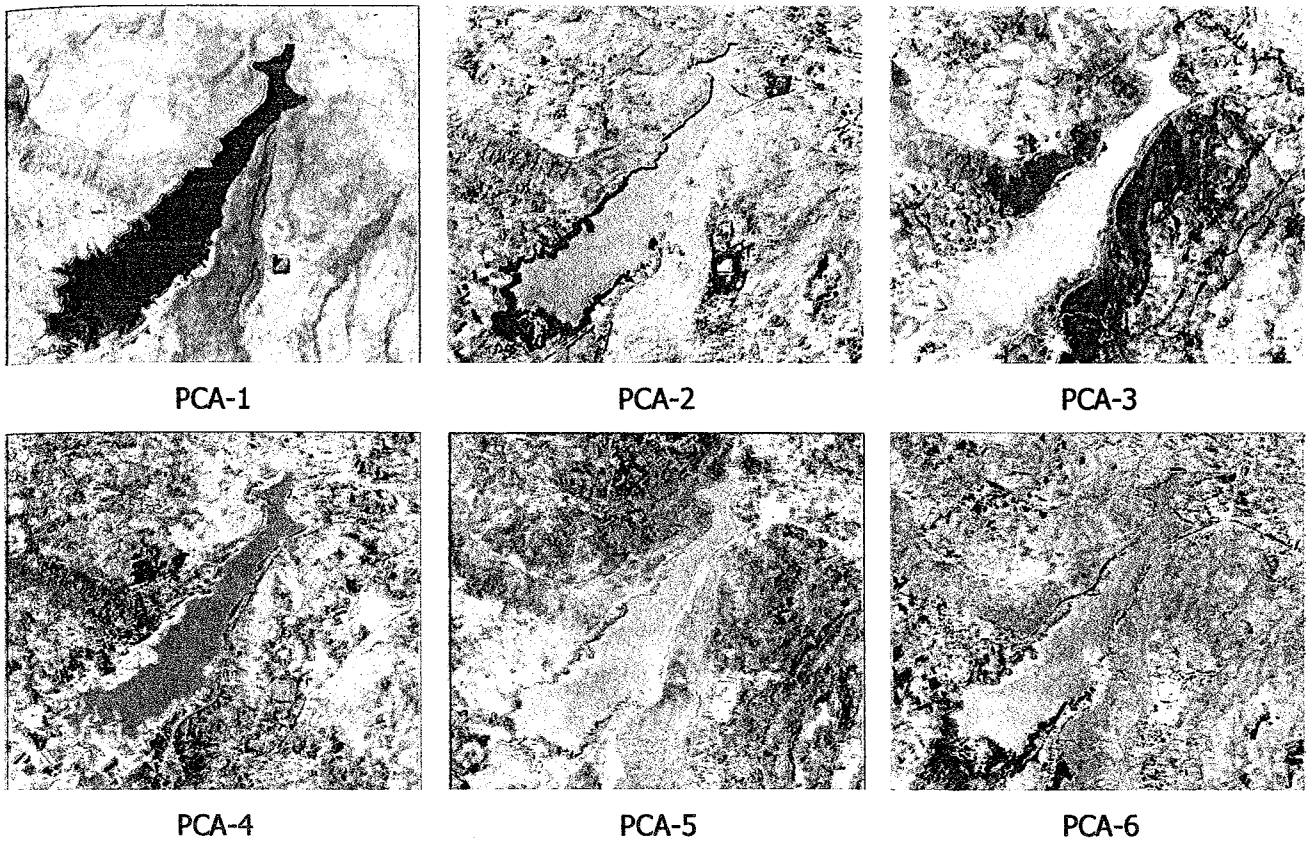


Figure 9-8: Principal components derived from a multiple-date dataset of Landsat TM in 1999 and 2004 of Lam Takhong Dam, Nakhon Ratchasima, Thailand. Principal component 2 contains change information about water surface in dam.

9.4.3 Image Algebra Change Detection

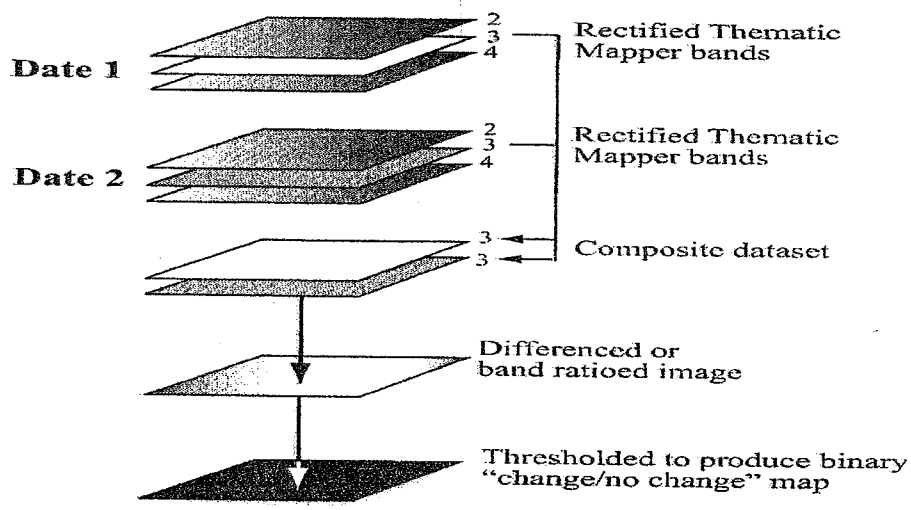
It is possible to identify the amount of change between two rectified images by band ratioing or image differencing. Image differencing involves subtracting the imagery of one date from that of another (Figure 9-9). If the two images have almost identical radiometric characteristics, the subtraction results in positive and negative values in areas of radiance change and zero values in areas of no change. The results are stored in a new change image.

When 8-bit data are analyzed in this manner, the potential range of difference values found in the change image is -255 to 255. (Figure 9-10) The results can be transformed into positive values by adding a constant, c (e.g., 127). The operation is expressed as:

$$\Delta BV_{ijk} = BV_{ijk}(1) - BV_{ijk}(2) + c \qquad (9\text{-}1)$$

where:

ΔBV_{ijk} = change pixel value
$BV_{ijk}(1)$ = brightness value on date 1
$BV_{ijk}(2)$ = brightness value on date 2
c = a constant (e.g., 127)
i = line number
j = column number
k = a single band (e.g., Landsat TM Band 3)



Advantages	Disadvantages
* Normally does not require atmospheric correction	* No "from–to" change information available
* Efficient method of identifying pixels that have changed in brightness value between dates	* Requires careful selection of the "change/no change" threshold

Figure 9-9: Algorithm, advantage and disadvantage of image algebra change detection (From Jensen, 2005).

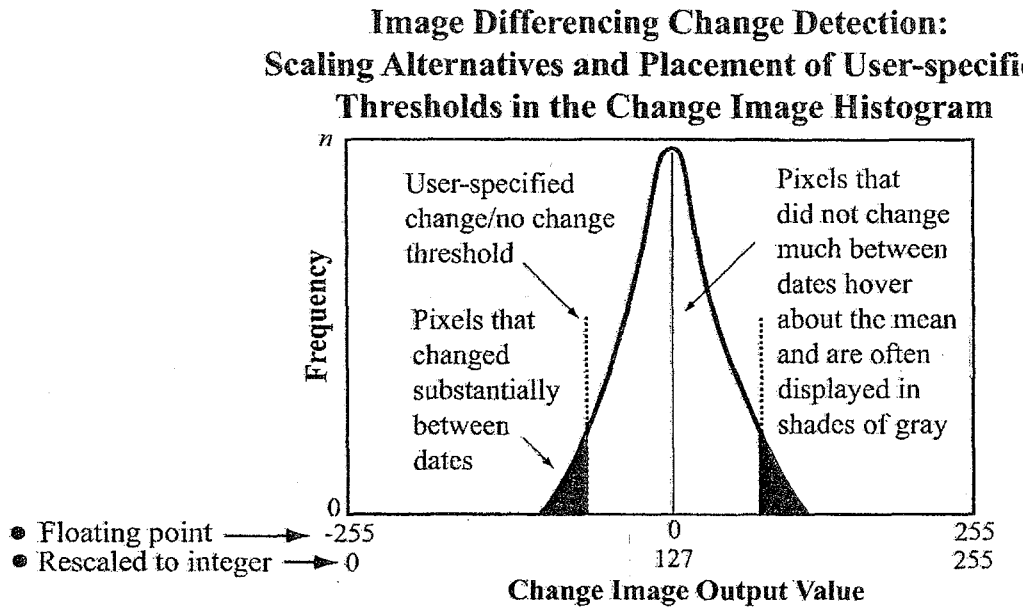


Figure 9-10: Image differencing change detection: Scaling alternatives and placement of user-specified thresholds in the change image histogram (From Jensen, 2005).

An example of image difference change detection is shown in Figure 9-11.

Differencing Vegetation Index Images

Image differencing does not have to be based on just the individual bands of remote sensor data. It may also be extended to comparing vegetation index information derived from multiple dates of imagery. For example, scientists have computed a normalized difference vegetation index (NDVI) on two dates and then subtracted one from another to determine change:

$$\Delta \text{NDVIij} = \text{NDVIij (1)} - \text{NDVIij (2)} + c \qquad (9\text{-}2)$$

where:

Δ NDVIij = change in NDVI value

NDVIij (1) = NDVI value on date 1

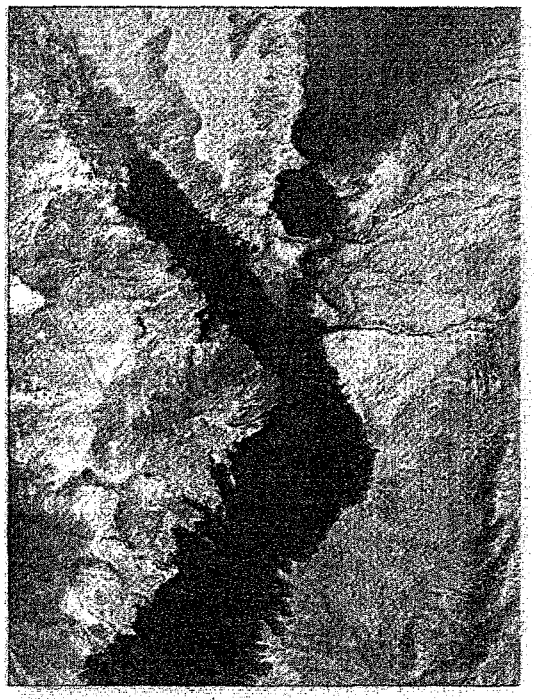
NDVIij (2) = NDVI value on date 2

i = line number

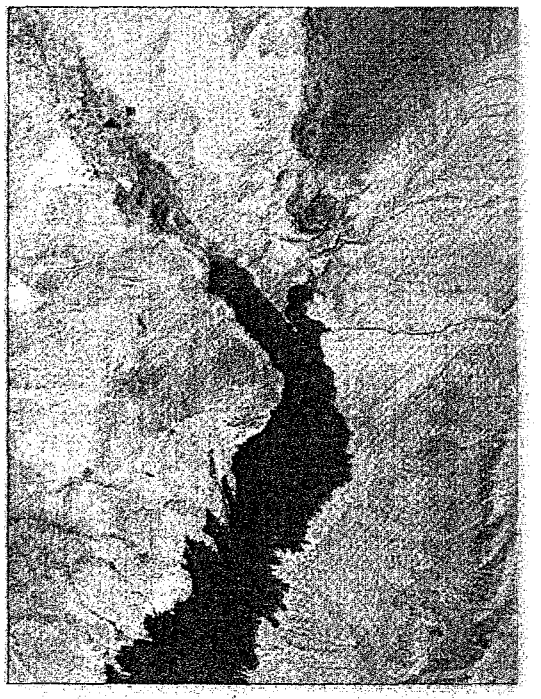
j = column number

c = a constant

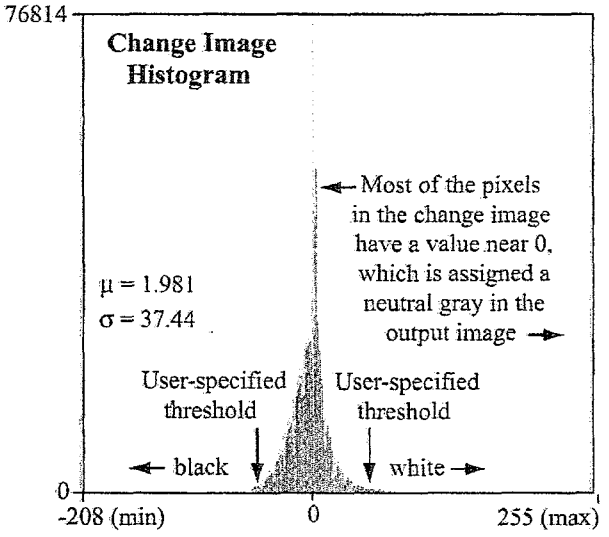
An example of image difference change detection using DNVI is shown in Figure 9-12.



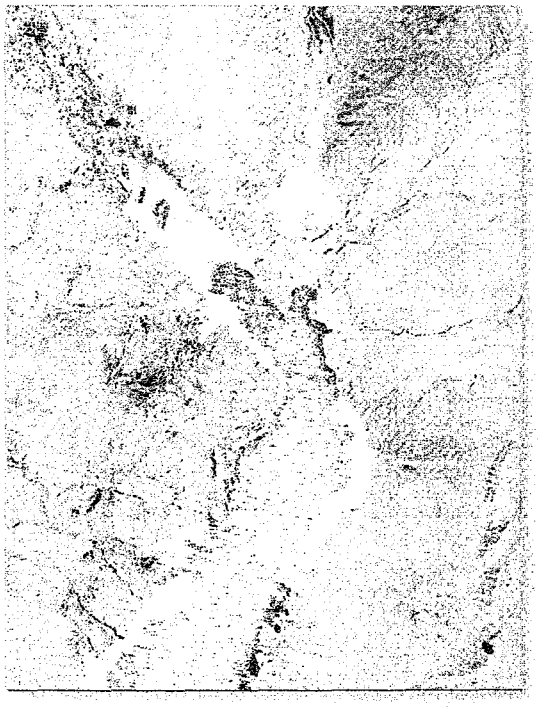
a. Landsat ETM$^+$ band 4 (NIR) May 3, 2000.



b. ASTER band 3 (NIR) April 19, 2003.

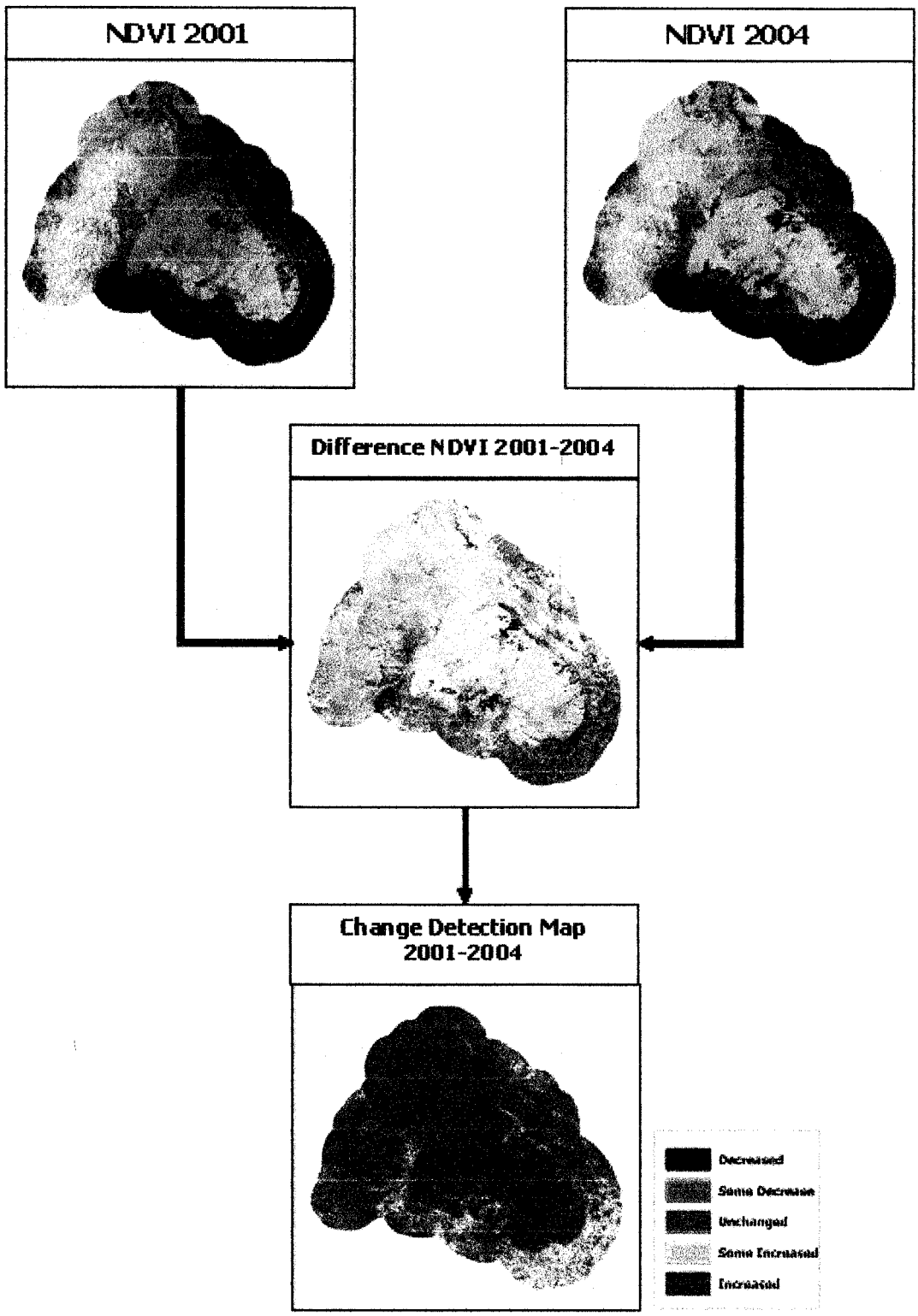


c. Histogram of the floating point image created by subtracting 2000 ETM$^+$ data from 2003 ASTER data. The symmetrical distribution confirms that the images were histogram-matched prior to processing.



d. Image differencing change detection based on Landsat ETM$^+$ band 4 May 3, 2000 data and ASTER band 3 April 19, 2003 data.

Figure 9-11: Image differencing change detection (from Jensen, 2005).



Vegetation Cover Change Detection between 2001 and 2004

Phu Kieo Wildlife Sanctuary, Chaiyaphum, Thailand.

Figure 9-12: Vegetation cover change detection based on differencing NDVI.

9.4.4 Post-classification Comparison Change Detection

Post-classification comparison change detection is a heavily used quantitative change detection method. It required rectification and classification of each remotely sensed image (Figure 9-13). The two maps are then compared on a pixel-by-pixel basis using a change detection matrix. Unfortunately, every error in the final change detection map will be present in the final change detection map. Therefore, it is imperative that the individual classification maps used in the post-classification change detection method be as accurate as possible. Post-classification comparison change detection is widely used and easy to understand. However, the accuracy of the change detection depends on the accuracy of the two separate classification maps. Example of post classification comparison change detection shows in Figure 9-14.

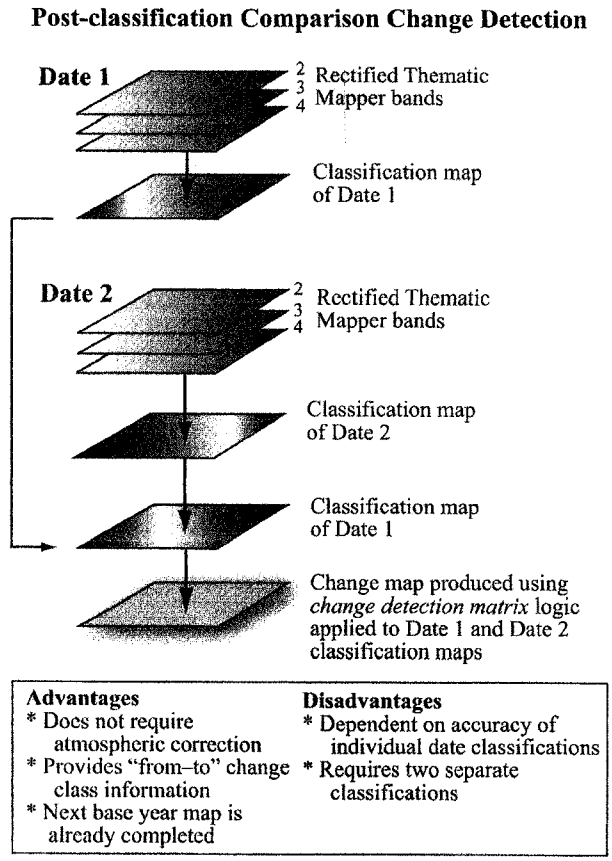


Figure 9-13: Algorithm, advantage and disadvantage of post-classification comparison change detection (Jensen, 2005).

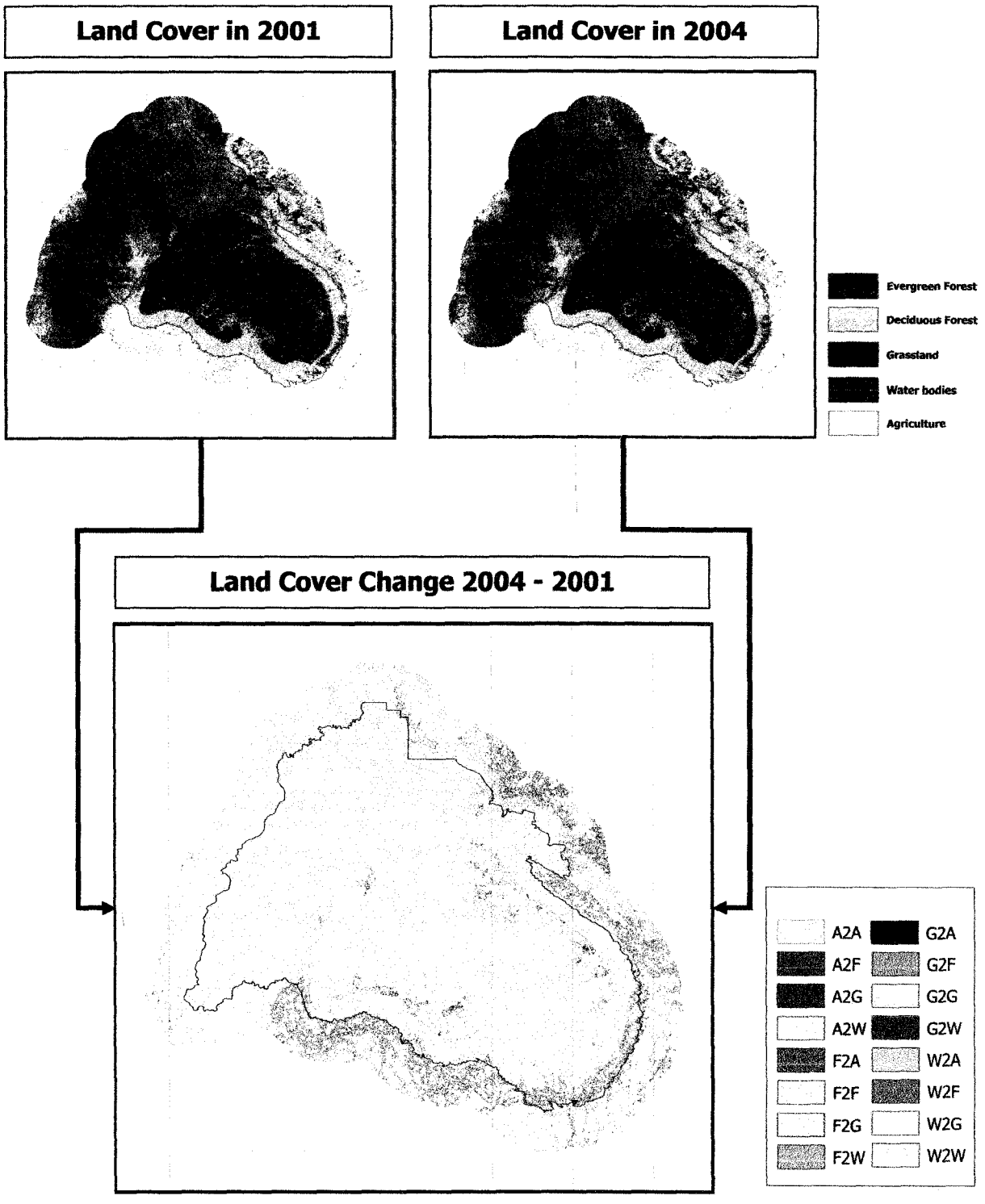


Figure 9-14: Land cover change 2001-2004, Phu Kieo Wildlife Sanctuary, Chaiyaphum, Thailand.

9.4.5 Change Detection Using a Binary Change Mask Applied to Date 2

This method of change detection is very effective. First, the analyst selected the basic image referred to as date 1 at time n. Date 2 may be earlier image (n-1) or later image (n+1). A traditional classification of date 1 is performed using rectified remote sensor data. Next, one of the bands (e.g., band 3) from both dates of imagery is placed in a new dataset. The two-band dataset is then analyzed using various image algebra change detection functions to produce a new change image file. The analyst usually selects a threshold value to identify areas of "change" and "no change" in the new image. Then, the change mask is overlaid onto Date 2 of the analysis and only those pixels that were detected as having changed are classified in the Date 2 imagery. Finally, a traditional post-classification comparison can then be applied to yield "from-to" change information (Figure 9-15). This method may reduce change detection errors and provides detailed "from-to" change information. The technique reduces effort by allowing analysts to focus on the small amount of area that changed between dates.

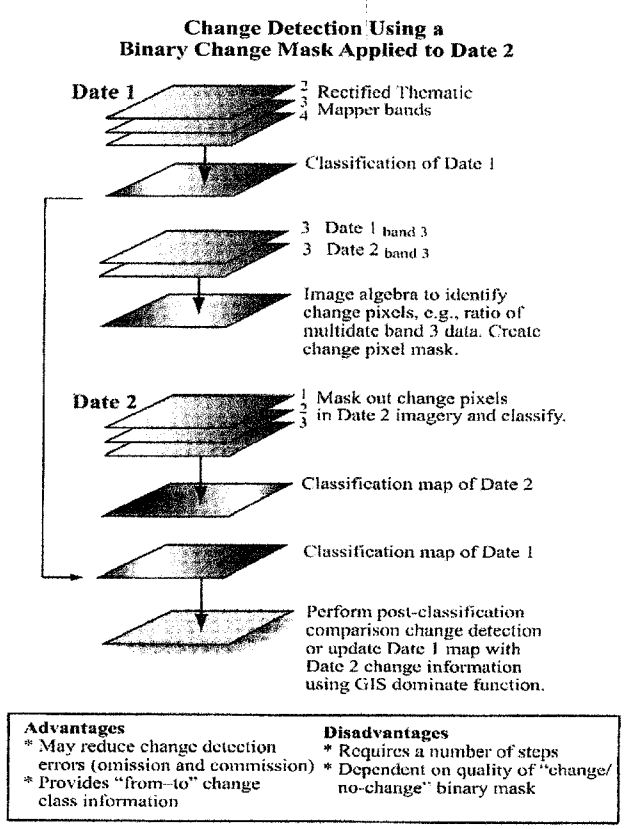


Figure 9-15: Algorithm, advantage and disadvantage of change detection using a binary change mask applied to date 2 (From Jensen, 2005).

9.4.6 Ancillary Data Source as Date 1

Sometimes there exists a land-cover data source that may be used in place of a traditional remote sensing image in the change detection process. Some of these data have been digitized. In stead of using a remotely sensed image as Date 1 in a coastal change detection project, it is possible to substitute the digital map of the region. In this case, the map is recorded to be compatible with classification scheme being used. Next, Date 2 of the analysis is classified and then compared on a pixel-by-pixel basis with the Date 1 information using post-classification comparison methods. Traditional "from-to" information can then be derived (Figure 9-16).

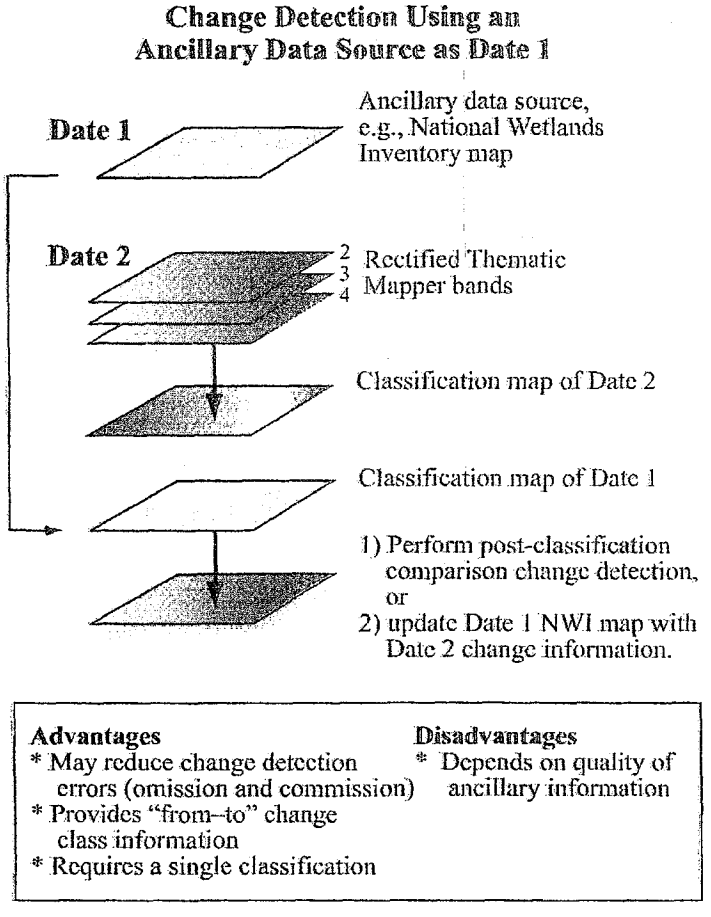


Figure 9-16: Algorithm, advantage and disadvantage of change detection using an ancillary as date 1 (From Jensen, 2005).

9.4.7 Spectral Change Vector Analysis

When land undergoes a change or disturbance between two dates, its spectral appearance normally changes. For example, consider the red and near-infrared spectral characteristics of a single pixel displayed in two-dimensional feature space (Figure 9-17). It appears that the land cover associated with this particular pixel has changed from Date 1 to Date 2 because the pixel resides at a substantially different location in the feature space on Date 2. The vector describing the direction and magnitude of change from Date 1 to Date 2 is spectral change vector (Melia, 1980; Chen et al., 2003). The total change magnitude per pixel (CMpixel) is computed by determining the Euclidean distance between end point through n-dimensional change space (Michalek, et al., 1993).

$$CM_{pixel} = \sum_{k=1}^{n} [BV_{ijk(date2)} - BV_{ijk(date1)}]^2 \quad (9\text{-}3)$$

where

BV i,j,k(date2) is the Date 1 pixel values in band k

BV i,j,k(date1) is the Date2 pixel values in band k.

A scale factor (e.g., 5) can be applied to each band to magnify small change in the data if desired. The change direction for each pixel is specified by whether the change is positive or negative in each band. Thus 2n possible types of changes can be determined per pixel (Virag and Cowell, 1987).

For example, if three bands are used there are 23 or 8 types of changes or sector code possible (see next Table). To demonstrate, let us consider a single registered pixel measured in three bands (1, 2 and 3) on two dates. If the change in band 1 was positive (e.g., BV i,j,1(date2)=45; BV i,j,1(date1) = 38; BVchange = 45-38=7), and the change in Band 2 was positive (e.g., BV i,j,2(date2)=20; BV i,j,2(date1) = 10; BVchange = 20-10=10) and the change in Band 3 was negative (e.g., BV i,j,3(date2)=25; BV i,j,3(date1) = 30; BVchange = 25-30=-5) then the change magnitude of the pixel would be CMpixel = 72 +102-52=174, and the change sector for this pixel would be "+,+,-" and have value of 7 as shown in the Table 9-1 and Figure 9-18.

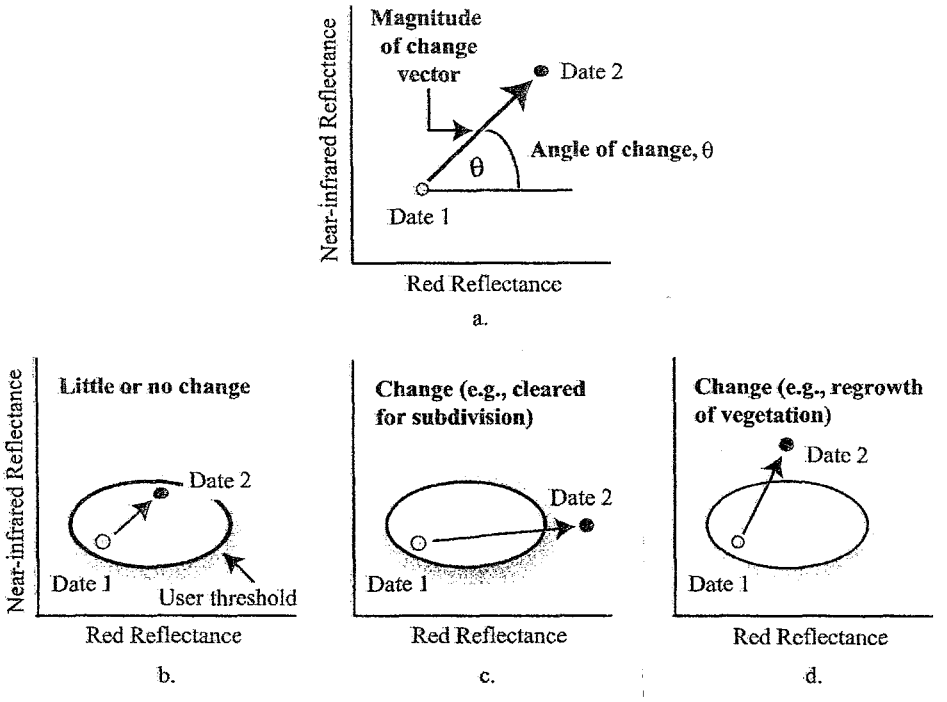


Figure 9-17: Schematic diagram of the spectral change detection method (From Jensen, 2005)

Table 9-1: Sector code definitions for change vector analysis using three bands

Sector Code	Change Detection		
	Band 1	**Band 2**	**Band 3**
1	-	-	-
2	-	-	+
3	-	+	-
4	-	+	+
5	+	-	-
6	+	-	+
7	+	+	-
8	+	+	+

Note: + indicate pixel value increase from Date 1 to Date 2
- indicate pixel value decrease from Date 1 to Date 2

From Jensen, 2005.

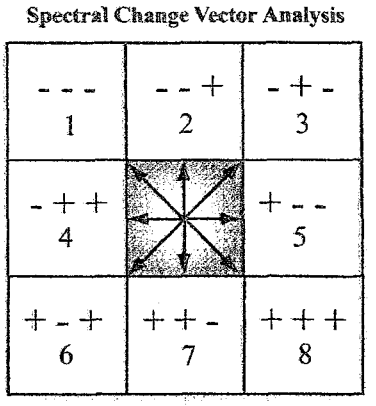


Figure 9-18: Possible change sector codes for a pixel measured in three bands on two dates.

Change vector analysis outputs two geometrically registered files; one contains the sector code and the other contains the sector vector magnitude. The change information may be superimposed onto an image of the study area with the change pixels color-code according to their sector code. This multispectral change magnitude image incorporates both the change magnitude and direction information. The decision that a change has occurred is made if a threshold is exceeded (Virag and Colwell, 1987). The threshold may be selected by examining the deep water area (if present), which should be unchanged, and recording their scaled magnitude from the change vector file. Figure 9-18b illustrated a case in which no change would be detected because threshold is not exceeded. Conversely, change would be detected in Figure 9-18c and 9-18d because the threshold was exceeded. The other half of the information contained in the change vector, that is, its direction. Direction contains information about the type of change.

9.4.8 Chi-square Transformation Change Detection

Ridd and Lui (1998) introduced a chi-square transformation change detection algorithm. It will work on any type of imagery. The chi-square transformation is:

$$Y_{pixel} = (X - M)^{-T} \sum^{-1} (X - M) \qquad (9\text{-}4)$$

where:

Ypixel	is the digital value of the pixel in the output change image;
X	is the vector of the difference of the six digital values between the two dates for each pixel;
M	is the vector of the mean residuals of each band for the entire image;
T	is transverse of the matrix
Σ-1	is the inverse covariance matrix of the six bands between the two dates.

The usefulness of the transformation in this context rests on the fact that Y is distributed as a chi-square random variable with p degrees of freedom where p is the number of bands. Y = 0 represents a pixel of no change. The user creates the output change image and highlights pixels with varying amounts of Y.

9.4.9 Cross-correlation Change Detection

Cross-correlation change detection makes use of an existing Date 1 land-cover map and a Date 2 unclassified multispectral dataset (Koeln and Bissonnette, 2000). First, every pixel in Date 1 land-cover map associated with a particular class c (e.g., forest) out of m possible classes is located in the Date 2 multispectral dataset (Figure 9-19). The mean (μck) and standard deviation (σ_{ck}) of all the brightness values in each band k in the Date 2 multispectral dataset associated with class c (e.g., forest) in Date 1 land-cover map are computed. Next, every pixel (BV_{ijk}) in the Date 2 scene associated with class c is compared with the mean (μ_{ck}) and divided by the standard deviation (σ_{ck}). This value is summed and squared over all k bands. This is performed for each class. The result is a Z-score associated with each pixel in the scene:

$$Z_{ijc} = \sum_{k=1}^{n} \left(\frac{BV_{ijk} - \mu_{ck}}{\sigma_{ck}} \right)^2 \qquad (9\text{-}5)$$

where:

Z_{ijc} is the Z-score for a pixel at location i,j in the Date 2 multispectral dataset associated with a particular class c found in the Date 1 land-cover map;

c is the Date 1 land-cover class under investigation;

n is the number of bands in the Date 2 multispectral image;

k is the band number in the Date 2 multispectral image;

BV_{ijk} is the brightness value of a pixel (or reflectance) at location i,j in band k of Date 2 multispectral dataset associated with a particular class found in the Date 1 land-cover map;

μ_{ck} is the mean of all pixels brightness values found in band k of Date 2 multispectral dataset associated with a particular class c in the Date 1 land-cover map;

σ_{ck} is the standard deviation of all pixels brightness values found in band k of Date 2 multispectral dataset associated with a particular class c in the Date 1 land-cover map.

The Z-statistic describes how close a pixel's response is to the expected spectral response of its corresponding class value in the land-cover map (Civco et al., 2002). In the output file, the

greater the z-score of an individual pixel, the greater the probability that its land cover has changed from Date 1 to Date 2. If desired, the image analyst can examine the Z-score image file and select a threshold that can be used to identify all pixels in the scene that have changed from Date 1 to Date 2. This information can be used to prepare a "change/no change' map of the area.

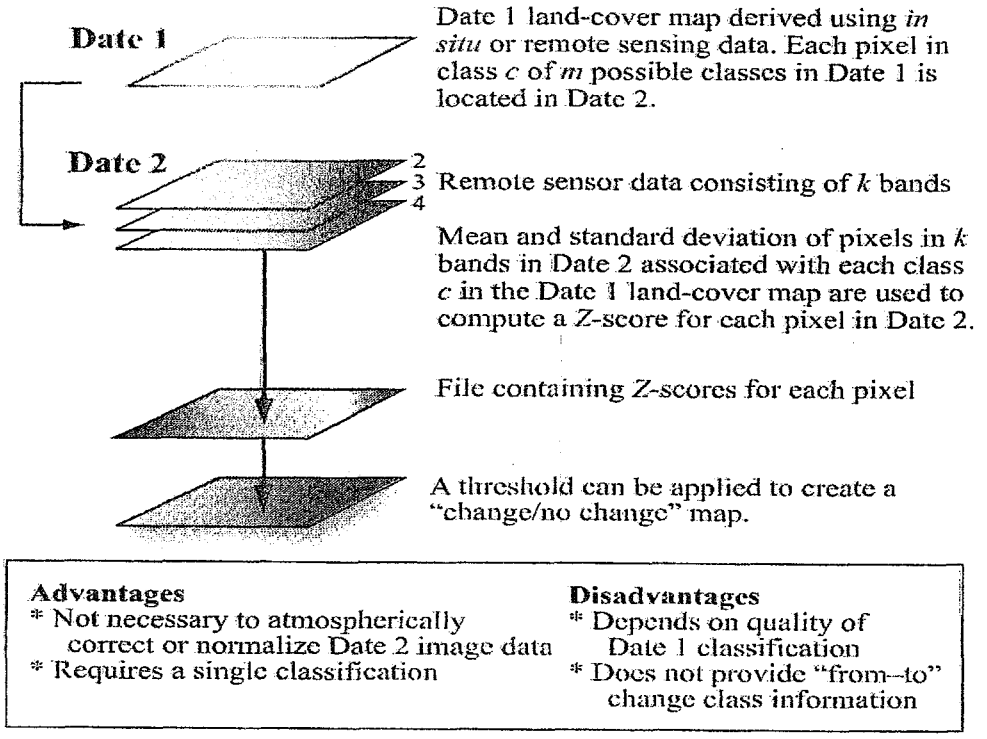


Figure 9-19: Algorithm, advantage and disadvantage of cross-correlation change detection (From Jensen, 2005).

9.4.10 Knowledge-based Vision Systems for Detecting Change

The use of expert systems to detect change automatically in an image with very little human interaction is still in its infancy. In fact, most scientists attempting to develop such systems have significant human intervention and employ many of the aforementioned change detection algorithms in the creation of knowledge-based change detection vision system. For example, Wang (1993) used a preprocessor to (1) perform image differencing, (2) create a change mask using PCA, (3) perform automated fuzzy supervised classification, and (4) extract attributes. Possible urban changes areas were then passed to a rule-based interpreter, which produced a change image.

9.4.11 Visual On-screen Detection and Digitization

A considerable amount of high-resolution remote sensor data is now available. Also, the aerial photography data are scanned (digitized) at high resolutions into digital image files. These dataset can then be registered to common base map and compared to identify change. Digitized high-resolution aerial photograph displayed on a CRT screen can be easily interpreted using standard photo interpretation techniques and the fundamental elements of image interpretation including size, shape, shadow, texture, etc (Jensen, 2000). Therefore, it is becoming increasingly common for analysts to visually interpret both dates of aerial photographs or very high resolution satellite data using heads-up on screen digitization and compare the various images to detect change as shown in Figure 9-20.

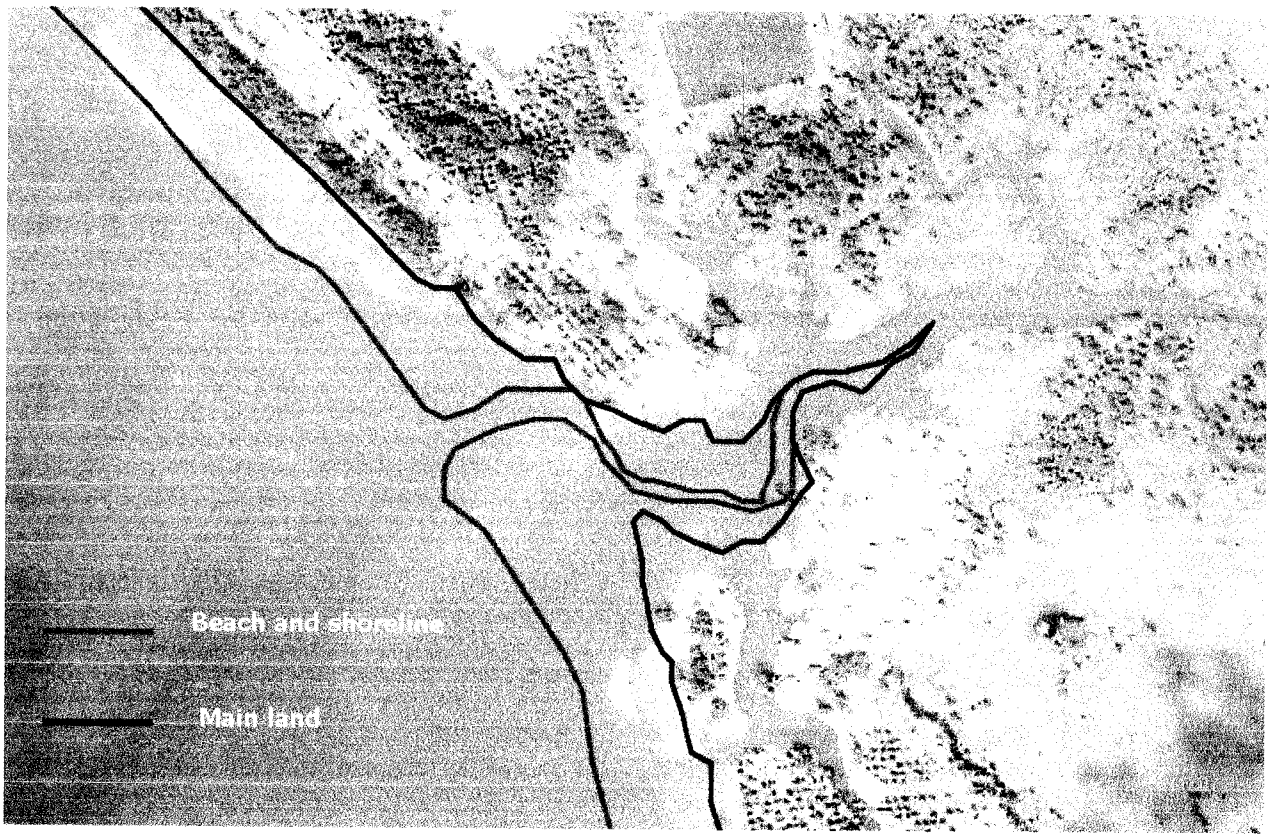


Figure 9-20: Visual on-screen change detection in Tsunami's effect at Ta Kau Pa District, Phang Nga, Thailand: Date 26 December 2004

REFERRENCE

Chapter 1: Concepts and Fundamentals of Remote Sensing

Barrett, E. C. and L. F. Curtis. 1982. Introduction to Environmental Remote Sensing. 2nd Edition. Chapman and Hall, London: 352 pp.

Campbell, J. B. 1987. Introduction to Remote Sensing. The Guilford Press, New York: 551 p.

Colwell, J. E. 1974. Vegetation Canopy Reflectance. Remote Sensing of Environment, 3: 175-183

Colwell, R. N. 1979. Remote sensing of natural resource-retrospect and prospect; Proceedings of Remote Sensing for Natural Resources, an International View of Problems, Promise and Accomplishments, University of Idao, Idao: pp. 48-68.

Colwell, R. N. 1983. Glossary. In Manual of Remote Sensing, 2 nd Edition (Editor R. N. Colwell) American Society of Photogrammetry, 1183-1207.

Colwell, R. N. 1997. History and Place of Photogrammetric Interpretation. In Manual of Photographic Interpretation, 2nd Ed., W. R. Phillipson (ed.), American Society for Photogrammetry & Remote Sensing, Virginia: 33-48.

Curran, P. J. 1985. Principle of Remote Sensing. Longman, London. 282 pp.

Elachi, C. 1987. Introduction to the Physics and Techniques of Remote Sensing. John Wiley & Sons, New York. 413. p

Heute, A. 2004. Remote Sensing of Soils and Soil Process. In Remote Sensing for Natural Resources Management and Environmental Monitoring: Manual of Remote Sensing 3rd , Vol., 4. Edited by Susan L, Ustin. John Wiley & Sons, Inc. NJ: pp 3-52.

Hoffer, R. M. 1978. Biological and Physical Considerations in Applying Computer-Aided Analysis Techniques to Remote Sensing Data. In Remote Sensing: The Quantitative Approach (Editors: P. H. Swain and S. M. Davis) MrGraw-Hill, New York: pp 227 -289.

Jensen, J. R. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd Edition. Practice Hall. 526 p.

Jensen, J. R. 2007. Remote Sensing of the Environment: An Earth Resource Perspective. 2nd Edition. Practice Hall. 592 p.

De Jong, S. M., van der Meer, F. D. and Jan G.P.W. Clevers. 2004. Basic of Remote Sensing. In Remote Sensing Image Analysis: Including the Spatial Domain. De Jong, S.M., van der Meer, F.D. (eds.) Kluwer Academic Publishers, the Netherlands. pp: 1–15.

Knipling, E. B. 1970. Physical and Physiological Basis for the reflectance of Visible and Near Infrared Radiation from Vegetation. Remote Sensing of Environment 1: 155-159.

Landgrebe, D. A. 2003. Signal Theory Methods in Multispectral Remote Sensing. John Wiley & Sons, New Jeysey. 508 p.

Lillesand, T. M. and R. W. Kiefer, 1979. Remote Sensing and Image Interpretation. John Wiley & Sons, Inc. New York, 612 pp.

Lillesand, T. M. and R. W. Kiefer, and J. W. Chipman. 2004. Remote Sensing and Image Interpretation. John Wiley & Sons, Inc. New York, 763 pp.

Lintz, J. Jr and D. S. Simonett, 1976. Preface and Introduction . In Remote Sensing of Environment. Editors J. Lintz, Jr. and D. S. Simonett. Addison-Wesley Publishing Company, Inc., London: 1-9 pp.

McCloy, K. R. 2006. Resources Management Information Systems: Remote Sensing, GIS and Modelling. 2Nd Edition. CRC Press Taylor & Francis Group, Fl, 575.

Sabins, F. F. 1987. Remote Sensing: Principle and Interpretation. 2nd Edition. W. H. Freeman and Company. New York. 449 p.

Simonett, D.S. 1983. The development and principles of remote sensing. In Manual of Remote Sensing, 2 nd Edition (Editor R. N. Colwell) American Society of Photogrammetry, 1-35.

Suits, J. M. 1983. Matter-Energy Interaction in the Optical Region. In Manual of Remote Sensing, 2 nd Edition (Editor R. N. Colwell) American Society of Photogrammetry, 61 - 113.

Tucker, C. J. 1987. An evaluation of the first four Landsat Thematic Mapper reflectance sensors for monitoring vegetation: A comparison with other satellite sensors systems. NASA Technical Memo, NASA TM-79617.

Chapter 2: Sensor in Remote Sensing

Jensen, J. R. 2007. Remote Sensing of the Environment: An Earth Resource Perspective. 2nd Edition. Practice Hall. 592 p.

Sabins, F. F. 1987. Remote Sensing: Principle and Interpretation. 2 nd Edition. W. H. Freeman and Company. New York. 449 p.

Chapter 3: Aerial Photography and Visual Interpretation

Campbell, J. B. 1987. Introduction to Remote Sensing. The Guilford Press, New York: 551 p.

Estes, J. E., E. J. Hajic and L. R. Tinney. 1983. Fundamentals of image analysis: Analysis of visible and thermal infrared data. In Manual of Remote Sensing, 2 nd Edition (Editor R. N. Colwell) American Society of Photogrammetry, 987-1124.

Jensen, J. R. 2007. Remote Sensing of the Environment: An Earth Resource Perspective. 2nd Edition. Practice Hall. 592 p.

Sabins, F. F. 1987. Remote Sensing: Principle and Interpretation. 2 nd Edition. W. H. Freeman and Company. New York. 449 p.

Chapter 4: Satellite Data and Digital Image Processing

Bernstein, R., 1983. Image Geometry and Rectification. In Manual of Remote Sensing, 2 nd Edition (Editor R. N. Colwell) American Society of Photogrammetry, 873-922.

Campbell, J. B. 1987. Introduction to Remote Sensing. The Guilford Press, New York: 551 p.

ERDAS. 2002. ERDAS Field Guide. Atlanta, Georgia. 658 p.

Green, W. B.,1983. Digital Image Processing: A Systems Approach. Van Nostrand Rheinhold Electrical/Computer Science and Engineering Series. Van Nostrand Reinhold Company. Ney York. 192 p.

Haralick, R. M. 1979. Statistical and structural approaches to texture. Proceedinds I.E.E.E.,67: 786-804.

Jensen, J. R. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd Edition. Practice Hall. 526 p.

Jensen, J. R. 2007. Remote Sensing of the Environment: An Earth Resource Perspective. 2nd Edition. Practice Hall. 592 p.

Landgrebe, D. A. 2003. Signal Theory Methods in Multispectral Remote Sensing. John Wiley & Sons, New Jeysey. 508 p.

Lillesand, T. M. and R. W. Kiefer, and J. W. Chipman. 2004. Remote Sensing and Image Interpretation. John Wiley & Sons, Inc. New York, 763 pp.

Maxwell, E. L., 1976. Multivariate system analysis of mutispectral imagery. Photogrammetric Engineering & Remote Sensing, 42( ):1173-1186.

Schowengerdt, R. A. 1983. Techniques for Image Processing and Classification. Academic Press, Inc. New York. 249 p.

Townshend, J. and C. Justice, 1981. Information Extraction from Remotely Sensed Data: A User View. Int. J. Remote Sensing, 2(4): 313-329.

Chapter 5: Preprocessing in Digital Image Processing

Bernstein, R., 1983. Image Geometry and Rectification. In Manual of Remote Sensing, 2 nd Edition (Editor R. N. Colwell) American Society of Photogrammetry, 873-922.

Campbell, J. B. 1987. Introduction to Remote Sensing. The Guilford Press, New York: 551 p.

Chavez, P. S., Jr. 1988. An Improved Dark-Object Subtraction Technique for Atmospheric Scattering Correction of Multispectral Data. Remote Sensing of Environment. 24: 459-479. Reference by Schowengerdt (1997)

Chavez, P. S., Jr. 1989. Radiometric Calibration of Landsat Thematic Mapper Multispectral Image. Photogrammetric Engineering & Remote Sensing, 55(9): 1285-1294. Reference by Schowengerdt (1997)

Chavez, P. S., Jr. 1996. Image-Based Atmospheric Correction-Revised and Improved Photogrammetric Engineering & Remote Sensing, 62(9): 1025-1036. Reference by Schowengerdt (1997)

ERDAS. 2002. ERDAS Field Guide. Atlanta, Georgia. 658 p.

Green, R. O. 2003. Introduction to Atmospheric Correction. Chapter 2 in ACORN Tutorial, Boulder, CO: Analytical Imaging and Geophysics, LLC. 12-18. Reference by Jensen (2005)

Holben, B., E. Vermote, Y. J. Kaufmann, D. Tanre, and V. Kalb. 1992. Aerosol Retrieval Over Land from AVHRR data-Application for Atmospheric Correction. IEEE Transaction on Geoscience and Remote Sensing. 30(2): 212-222. Reference by Schowengerdt (1997)

Jensen, J. R. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd Edition. Practice Hall. 526 p.

Lavreau, J. 1991. De-hazing Landsat Thematic Mapper Images. Photogrammetric Engineering & Remote Sensing, 57(10): 1297-1302. Reference by Schowengerdt (1997)

Lillesand, T. M. and R. W. Kiefer, and J. W. Chipman. 2004. Remote Sensing and Image Interpretation. John Wiley & Sons, Inc. New York, 763 pp.

Meyers, P., Itten, K. L., Kellenberger, T., Sandmeier, S. And R. Sandmeier. 1993. Radiometric Corrections of Topographically Induced Effects on Landsat TM Data in an Alpine Environment. ISPRS Journal of Photogrammetry and Remote Sensing 48(4): 17-28. Reference by Jensen (2005)

Moran, M. S., R. D. Jackson, P. N. Slater, and P. M. Teillet. 1992. Evalation of Simplified Procedures for Retrieval of Land Surface Reflectance Factors from Satellite Sensor Output. Remote Sensing of Environment. 41: 169-184. Reference by Schowengerdt (1997)

Potter, J. F. 1984. The Channel Correlation Method for Estimating Aerosol Levels from Multispectral Scanner Data. Photogrammetric Engineering & Remote Sensing, 50: 43-52. Reference by Schowengerdt (1997)

Potter, J. F. and M. Mendolowitz. 1975. On the Determination of the Haze Levels from Landsat Data. In 10th International Sysposium on Remote Sensing of Environment, Ann Arbor, MI, Environmental Research Institute of Michigan,: 695-703. Reference by Schowengerdt (1997)

Richter, R. 1996a. A Spatilally Adaptive Fast Atmospheric Correction Algorithm. International Journal of Remote Sensing. 17(6): 1201-1214. Reference by Schowengerdt (1997)

Richter, R. 1996b. Atmospheric Correction of Satellite Data with Haze Removal Including a Haze/Clear Transition Region. Computers & Geosciences 22(6):675-681. Reference by Schowengerdt (1997)

Schowengerdt, R. A. 1997. Remote Sensing: Models and Methods for Image Processing. Academic Press, Inc. New York. 522 p.

Singh, S. M. and A. P. Cracknell. 1986. The Estimation of Atmospheric Effects for SPOT using AVHRR Channel-1 Data. International Journal of Remote Sensing. 7(3): 361-377. Reference by Schowengerdt (1997)

Switzer, P., W.S. Kowalik, and R.J. Lyon. 1981. Estimation of Atmospheric Path Radiance by Covariance Matrix Method. Photogrammetric Engineering & Remote Sensing, 47: 1469-1476. Reference by Schowengerdt (1997)

Teillet, P.M., Guindon, B. and D.G. Goodenough. 1982. On the Slope-aspect Correction of Multispectral Scanner Data. Canadian Journal of Remote Sensing. 8(2):84-106. Reference by Jensen (2005)

Teilet, P. M. and G. Fedosejevs. 1995. On the Dark Target Approach to Atmospheric Correction of Remotely Sensed Data. Canadian Journal of Remote Sensing 21(4): 374-387. Reference by Schowengerdt (1997)

Wrigley, R.C., M.A. Spanner, R.E. Slye, R.F. Pueschel, and H.R. Aggarwal. 1992. Atmospheric Correction of Remotely Sensed Image Data by a Simpligied Model. Journal of Geophysical Research 97(D17): 18797-18814. Reference by Schowengerdt (1997)

Chapter 6: Image Enhancement in Digital Image Processing

Chavez, P. S., Jr., S. C. Sides, and J. A. Anderson. 1991. Comparison of Three Different Methods to Merge Multiresolution and Multispectral Data: Landsat TM and SPOT Panchromatic. Photogrammetric Engineering & Remote Sensing 57 (3): 295-303. (Reference by ERDAS (2002)

ERDAS. 2002. ERDAS Field Guide. Atlanta, Georgia. 658 p.

Fahnestock, J. D., and R. A. Schowengerdt. 1983. Spatially Variant Contrast Enhancement Using Local Range Modification.Optical Engineering 22 (3): 378-381. (Reference by ERDAS (2002)

Faust, N. L. 1989. Image Enhancement. Volume 20, Supplement 5 of Encyclopedia of Computer Science and Technology. Ed. A. Kent and J. G. Williams. New York: Marcel Dekker, Inc.

Franklin, S.E., Maudie, A.J. and M.B. Lavigne, 2001. "Using Spatial Co-occurrence Texture to increase Forest Structure and Species Composition Classification Accuracy". Photogrammetric Engineering & Remote Sensing, 67(7):849-855. (Reference by Jensen, 2005)

Jensen, J. R. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd Edition. Practice Hall. 526 p.

Lillesand, T. M. and R. W. Kiefer, and J. W. Chipman. 2004. Remote Sensing and Image Interpretation. John Wiley & Sons, Inc. New York, 763 pp.

McCloy, K. R. 2006. Resources Management Information Systems: Remote Sensing, GIS and Modelling. 2Nd Edition. CRC Press Taylor & Francis Group, Fl, 575.

Peli, T., and J. S. Lim. 1982. Adaptive Filtering for Image Enhancement. Optical Engineering 21 (1): 108-112. (Reference by ERDAS (2002)

Crippen, R. E. 1989. A Simple Spatial Filtering Routine for the Cosmetic Removal of Scan-Line Noise from Landsat TM P-Tape Imagery. Photogrammetric Engineering & Remote Sensing 55 (3): 327-331. (Reference by ERDAS (2002)

Schwartz, A. A., and J. M. Soha. 1977. Variable Threshold Zonal Filtering. Applied Optics 16 (7). (Reference by ERDAS (2002)

Schowengerdt, R. A. 1980. Reconstruction of Multispatial, Multispectral Image Data Using Spatial Frequency Content. Photogrammetric Engineering & Remote Sensing 46 (10): 1325-1334. (Reference by ERDAS (2002)

Schowengerdt, R. A. 1997. Remote Sensing: Models and Methods for Image Processing. Academic Press, Inc. New York. 522 p.

Welch, R., and W. Ehlers. 1987. Merging Multiresolution SPOT HRV and Landsat TM Data. Photogrammetric Engineering & Remote Sensing 53 (3): 301-303. (Reference by ERDAS (2002)

Chapter 7: Image Classification in Digital Image Processing

Bezdek, J.C., Ehrlich, R. & Full, W. (1984) FCM: Th e fuzzy c-means clustering algorithm, Computers and Geosciences 10: 191-203. (Reference by Foody 2004)

Duda, R. O., Hart, P. E. and D. G. Stork. Pattern Classification. John Wiley & Sons, USA. 654 p.

ERDAS. 2002. ERDAS Field Guide. Atlanta, Georgia. 658 p.

Foody, G. M. 2004. Sub-Pixel Methods in Remote Sensing in In Remote Sensing Image Analysis: Including the Spatial Domain. De Jong, S.M., van der Meer, F.D. (eds.) Kluwer Academic Publishers, the Netherlands. pp: 37–49.

Jensen, J. R. and D. C. Toll, 1982. "Detecting Residential Land Use Development at the Urban Fringe." Photogrammetric Engineering & Remote Sensing, 48:629-643. (Reference by Jensen, 2005)

Jensen, J. R., Qiu, F. amd M. Ji, 1999. "Predictive Modeling of Coniferous Forest Age Using Statistcal and Artificial Neural Network Approaches Applied to Remote Sensing Data." International Journal of Remote Sensing, 20(14): 2805-2822. (Reference by Jensen, 2005)

Jensen, J. R. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd Edition. Practice Hall. 526 p.

Keiner, L. E. 1999. Neural Network as Non-linear Function Approximators for Remote Sensing Applications. In Information Processing for Remote Sensing (Ed. C. H. Chen) World Scientific , Singapore: 213-223.

Kloer, B. R. 1994. Hybrid Parametric/Non-Parametric Image Classification. Paper presented at the ACSM-ASPRS Annual Convention, Reno, Nevada, April 1994.

Krapivin, V. F. and G. W. Phillips, 2001. " A Remote Sensing-Based Expert Systems to Study the Aral-Caspian Aquageosystem Water Regime." Remote Sensing of Environment, 75: 201-215. (Reference by Jensen, 2005)

Lillesand, T. M. and R. W. Kiefer, and J. W. Chipman. 2004. Remote Sensing and Image Interpretation. John Wiley & Sons, Inc. New York, 763 pp.

Matsuyama, T. 1989. " Expert Systems for Image Processing: Knowledge-based Composition of Image Analysis Processes." Computer Vision, Graphics and Image Processing, 48(1):22-49. (Reference by Jensen, 2005)

McCloy, K. R. 2006. Resources Management Information Systems: Remote Sensing, GIS and Modelling. 2Nd Edition. CRC Press Taylor & Francis Group, Fl, 575.

Merriam-Webster, 2003. Merriam-Webster Dictionary, Springfield, MA. (Reference by Jensen, 2005)

Moller-Jensen, L. 1997. "Classification of Urban Land Cover Based on Expert Systems, Object Models and Texture." Computers, Environment and Urban Systems, 21(3/4):291-302. (Reference by Jensen, 2005)

Muchony, D., Borak, J., Chi, H., Friedl, M., Gopal, S., Hodges, J., Morrow, N. and A. H. Strahler, 2000. "Application of MODIS Global Supervised Classification Model to Vegetation and land Cover Mapping in Central America." International Journal of Remote Sensing, 21:1115-1138. (Reference by Jensen, 2005)

Pal, M. and P. M. Mather. 2003. "An Assessment of the Effectiveness of Decision Tree Methods for Land Cover Clssification." Remote Sensing of Environment, 86:554-565. (Reference by Jensen, 2005)

Richards, J. A. and X. Jia. 1999. Remote Sensing Digital Image Analysis: An Introduction, Berlin Springer, 363 p. (Reference by Jensen, 2005)

Schowengerdt, R. A. 1997. Remote Sensing: Models and Methods for Image Processing. Academic Press, Inc. New York. 522 p.

Stefanov, W. L., Ramsey, M. S. and P. R. Christensen, 2001. "Monitoring Urban Land Cover Change: An Expert System Approach to Land Cover Classification of Semiarid to Arid Urban Centers." Remote Sensing of Environment, 77:173-185. (Reference by Jensen, 2005)

Stow, D., Coulter, L., Kaiser, J., Hope, A., Service, D., Schutte, K. and A. Walters, 2003, "Irrigated Vegetating Assessments for Urban Environments," Photogrammetric Engineering & Remote Sensing, 69(4):381-390. (Reference by Jensen, 2005)

Swain, P. H., and S. M. Davis. 1978. Remote Sensing: The Quantitative Approach. New York: McGraw Hill Book Company. 396 p.

Tso, B. and P.M. Mather, 2001, Classificarion Methods for Remotely Sensed Data, New York: Taylor & Francis, 322 p. (Reference by Jensen, 2005)

Werbos, P.J., 1974, Beyond Regression: New Tools for Prediction and Analysis in Behavioral Sciences, Ph.D. Thesis, Cambridge, Ma: Harvard University. (Reference by Jensen, 2005)

Yang, C. and P. Chung, 2002, "Knowledge-based Automatic Change Detection Positioning System for Complex Heterogeneous Environments" Journal of Intelligent and Robotic Systems, 33:85-98. (Reference by Jensen, 2005)

Chapter 8: Accuracy Assessment

Congalton, R. G. 1988. Using Spatial Autocorrelation Analysis to Explore the Errors in Maps Generated from Remotely Sensed Data" Photogrammetric Engineering & Remote Sensing, 54(5):587-592. (Reference by Jensen, 2005)

Congalton, R. G. 1991. " A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." Remote Sensing of Environment, 37:35-46. (Reference by Jensen, 2005)

Congalton, R. G. and K. Green, 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, Boca Raton, FL: Lewis Publishers, 137 p.

Fitzpatrick-Lins, K. 1981. Comparison of Sampling Procedures and Data Analysis for a Land-use and Land-cover Map. Photogrammetric Engineering & Remote Sensing, 55(4):475-478. (Reference by Jensen, 2005)

Jensen, J. R. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd Edition. Practice Hall. 526 p.

Landis, J. and G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. Biometrics. 33: 159-174. (Reference by Jensen, 2005)

King, L., 1969, Statistical Analysis in Geography, Englewood Cliffs, NJ: Prentice-Hall, 288 p. (Reference by Jensen, 2005)

Story, M. and R. Congalton, 1986, "Accuracy Assessment: A User's Perspective," Photogrammetric Engineering & Remote Sensing, 52(3):397-399. (Reference by Jensen, 2005)

Chapter 9: Digital Change Detection

Chen, J., Gong, P., He, C., Pu, R. and P. Shi, 2003. "Land-Use/Land-Cover Change Detection Using Improved Change-Vector Analysis." Photogrammetric Engineering & Remote Sensing 69(4):369-379. (Reference by Jensen, 2005)

Civco, D.L. 1989. "Topographic Normalization of Landsat Thematic Mapper Digital Imagery." Photogrammetric Engineering & Remote Sensing 55(9):1303-1309. (Reference by Jensen, 2005)

Civco, D. L., Hund, J. D., Wilson, E.H., Song, M. and Z. Zhang, 2002. "A Comparison of Land Use and Land Cover Change Detection Methods." Proceedings, ASPRS-ACSM Annual Conference and FIG XXII Congress, Bethesda, MD: American Society for Photogrammetry & Remote Sensing, 10 p. (Reference by Jensen, 2005)

Foody, G. M. 2001. "Monitoring the Magnitude of Land-Cover Change Around the Southern Limits of the Sahara." Photogrammetric Engineering & Remote Sensing 67(7):841-847. (Reference by Jensen, 2005)

Jensen, J. R. 1981. "Urban Change Detection Mapping Using Landsat Digital Data." The American Cartographer, 8(2):127-147. (Reference by Jensen, 2005)

Jensen, J. R., Cowen, D. J., Narumalani, S., Althausen, J. D. and O. Weatherbee. 1993. "Am Evaluation of Coast Watch Change Detection Protocol in South Carolina." Photogrammetric Engineering & Remote Sensing 59(6):1039-1046. (Reference by Jensen, 2005)

Jensen, J. R. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd Edition. Practice Hall. 526 p.

Koeln, G. and J. Bissonnette, 2000. "Cross-correction Analysis: Mapping Land Cover Change with a Historic Land Cover Database and a Recent, Single-date Multispectral Image." Proceedings, American Society for Photogrammetry & Remote Sensing, Bethoda, MD: ASPRS, 8 p. (Reference by Jensen, 2005)

Lunetta, R. S. 1999. Applications, Project Formulation, and Analytical Approach. In Remote Sensing Change Detection: Environmental Monitoring Methods and Applications (Editors Lunetta, R. S. and C. D. Elvidge), New York: Taylor &Francis, p 1-19.

Melia, W. A. 1980. "Change Vector Analysis: An Approach for Detecting Forest Changes with Landsat, " Proceedings, LARS Machine Processing of Remotely Sensed Data Sysmposium, W. Lafayette, In: Laboratory for the Applications of Remote Sensing, pp. 326-336. (Reference by Jensen, 2005)

Michalek, J. L., Wagner, T., Luczkovich, J. J. and R. W. Stoffle, 1993. "Multispectral Change Vector Analysis for Monitoring Coastal Marine Environments." Photogrammetric Engineering & Remote Sensing 59(3):381-384. (Reference by Jensen, 2005)

Ridd, M. K. and J. Liu, 1998. "A Comparison of Four Algorithms for Change Detection in an Urban Environment." Remote Sensing of Environment, 63:95-100. (Reference by Jensen, 2005)

Skole, D. 1994. "Data on Global Land-cover Change: Acquisition, Assessment and Analysis." in W. B. Meyer and B. L. Turner (Eds.), Change in Land Use and Land Cover: A Global Perspective, Cambridge: Cambridge University Press, 437-471. (Reference by Jensen, 2005)

Virag, L. A. and J. E. Colwell, 1987. "An Improved Procedure for Analysis of Change in Thematic Mapper-Pairs." Proceedings of the 21st International Symposium on Remote Sensing of Environment, Ann Arbor: Environmental Research Institute of Michigan, 1101-1110. (Reference by Jensen, 2005)

Wang, F. 1993. " A knowledge-based Vision System for Detecting Land Changes at Urban Fringes. IEEE Transactions on Geoscience & Remote Sensing, 31:136-145. (Reference by Jensen, 2005)

Yuan, D., Elvidge, C. D. and R. S. Lunetta. 1999. Survey of Multispectral Methods for Land Cover Change Analysis. In Remote Sensing Change Detection: Environmental Monitoring Methods and Applications (Editors Lunetta, R. L. and C. Elvidge), New York: Taylor &Francis, p 21-39. (Reference by Jensen, 2005)