

John A. Richards

Remote Sensing Digital Image Analysis

Sixth Edition

Contents

1	Sources and Characteristics of Remote Sensing Image Data	1
1.1	Energy Sources and Wavelength Ranges	1
1.2	Primary Data Characteristics	5
1.3	Remote Sensing Platforms	6
1.4	What Earth Surface Properties Are Measured?	11
1.4.1	Sensing in the Visible and Reflected Infrared Ranges	11
1.4.2	Sensing in the Thermal Infrared Range	14
1.4.3	Sensing in the Microwave Range	14
1.5	Spatial Data Sources in General and Geographic Information Systems	18
1.6	Scale in Digital Image Data	21
1.7	Digital Earth	21
1.8	How This Book Is Arranged	23
1.9	Bibliography on Sources and Characteristics of Remote Sensing Image Data	25
1.10	Problems	27
2	Correcting and Registering Images	31
2.1	Introduction	31
2.2	Sources of Radiometric Distortion	32
2.3	Instrumentation Errors	32
2.3.1	Sources of Distortion	32
2.3.2	Correcting Instrumentation Errors	34
2.4	Effect of the Solar Radiation Curve and the Atmosphere on Radiometry	35
2.5	Compensating for the Solar Radiation Curve	37
2.6	Influence of the Atmosphere	38
2.7	Effect of the Atmosphere on Remote Sensing Imagery	42
2.8	Correcting Atmospheric Effects in Broad Waveband Systems	43

2.9	Correcting Atmospheric Effects in Narrow Waveband Systems	45
2.10	Empirical, Data Driven Methods for Atmospheric Correction	49
2.10.1	Haze Removal by Dark Subtraction	50
2.10.2	The Flat Field Method	50
2.10.3	The Empirical Line Method	51
2.10.4	Log Residuals	52
2.11	Sources of Geometric Distortion	53
2.12	The Effect of Earth Rotation	54
2.13	The Effect of Variations in Platform Altitude, Attitude and Velocity	56
2.14	The Effect of Sensor Field of View: Panoramic Distortion	56
2.15	The Effect of Earth Curvature	59
2.16	Geometric Distortion Caused by Instrumentation Characteristics	60
2.16.1	Sensor Scan Nonlinearities	61
2.16.2	Finite Scan Time Distortion	61
2.16.3	Aspect Ratio Distortion	61
2.17	Correction of Geometric Distortion	62
2.18	Use of Mapping Functions for Image Correction	62
2.18.1	Mapping Polynomials and the Use of Ground Control Points	63
2.18.2	Building a Geometrically Correct Image	64
2.18.3	Resampling and the Need for Interpolation	65
2.18.4	The Choice of Control Points	67
2.18.5	Example of Registration to a Map Grid	68
2.19	Mathematical Representation and Correction of Geometric Distortion	70
2.19.1	Aspect Ratio Correction	70
2.19.2	Earth Rotation Skew Correction	71
2.19.3	Image Orientation to North–South	72
2.19.4	Correcting Panoramic Effects	72
2.19.5	Combining the Corrections	72
2.20	Image to Image Registration	73
2.20.1	Refining the Localisation of Control Points	73
2.20.2	Example of Image to Image Registration	75
2.21	Other Image Geometry Operations	78
2.21.1	Image Rotation	78
2.21.2	Scale Changing and Zooming	78
2.22	Bibliography on Correcting and Registering Images	79
2.23	Problems	80

3	Interpreting Images	87
3.1	Introduction	87
3.2	Photointerpretation	88
3.2.1	Forms of Imagery for Photointerpretation	89
3.2.2	Computer Enhancement of Imagery for Photointerpretation	90
3.3	Quantitative Analysis: From Data to Labels	91
3.4	Comparing Quantitative Analysis and Photointerpretation	93
3.5	The Fundamentals of Quantitative Analysis	94
3.5.1	Pixel Vectors and Spectral Space	94
3.5.2	Linear Classifiers	98
3.5.3	Statistical Classifiers	98
3.6	Sub-classes and Spectral Classes	101
3.7	Unsupervised Classification	103
3.8	Bibliography on Interpreting Images	103
3.9	Problems	104
4	Radiometric Enhancement of Images	107
4.1	Introduction	107
4.1.1	Point Operations and Look Up Tables	107
4.1.2	Scalar and Vector Images	108
4.2	The Image Histogram	108
4.3	Contrast Modification	109
4.3.1	Histogram Modification Rule	109
4.3.2	Linear Contrast Modification	110
4.3.3	Saturating Linear Contrast Enhancement	111
4.3.4	Automatic Contrast Enhancement	112
4.3.5	Logarithmic and Exponential Contrast Enhancement	113
4.3.6	Piecewise Linear Contrast Modification	113
4.4	Histogram Equalisation	113
4.4.1	Use of the Cumulative Histogram	113
4.4.2	Anomalies in Histogram Equalisation	120
4.5	Histogram Matching	122
4.5.1	Principle	122
4.5.2	Image to Image Contrast Matching	123
4.5.3	Matching to a Mathematical Reference	126
4.6	Density Slicing	126
4.6.1	Black and White Density Slicing	126
4.6.2	Colour Density Slicing and Pseudocolouring	127
4.7	Bibliography on Radiometric Enhancement of Images	129
4.8	Problems	131

5	Geometric Processing and Enhancement: Image Domain Techniques	135
5.1	Introduction	135
5.2	Neighbourhood Operations in Image Filtering	136
5.3	Image Smoothing	138
5.3.1	Mean Value Smoothing	138
5.3.2	Median Filtering	139
5.3.3	Modal Filtering	140
5.4	Sharpening and Edge Detection	140
5.4.1	Spatial Gradient Methods	141
5.4.1.1	The Roberts Operator	143
5.4.1.2	The Sobel Operator	143
5.4.1.3	The Prewitt Operator	144
5.4.1.4	The Laplacian Operator	145
5.4.2	Subtractive Smoothing (Unsharp Masking)	147
5.5	Edge Detection	147
5.6	Line and Spot Detection	150
5.7	Thinning and Linking	150
5.8	Geometric Processing as a Convolution Operation	151
5.9	Image Domain Techniques Compared with Using the Fourier Transform	153
5.10	Geometric Properties of Images	154
5.10.1	Measuring Geometric Properties	155
5.10.2	Describing Texture	156
5.11	Morphological Analysis	159
5.11.1	Erosion	161
5.11.2	Dilation	162
5.11.3	Opening and Closing	163
5.11.4	Boundary Extraction	164
5.11.5	Other Morphological Operations and Applications	166
5.12	Object and Shape Recognition	166
5.13	Bibliography on Geometric Processing and Enhancement: Image Domain Techniques	167
5.14	Problems	168
6	Spectral Domain Image Transforms	171
6.1	Introduction	171
6.2	Image Arithmetic and Vegetation Indices	172
6.3	The Principal Components Transform	174
6.3.1	The Mean Vector and the Covariance Matrix	174
6.3.2	A Zero Correlation, Rotational Transform	179
6.3.3	The Effect of an Origin Shift	184
6.3.4	Example and Some Practical Considerations	185

6.3.5	Application of Principal Components in Image Enhancement and Display	187
6.3.6	The Taylor Method of Contrast Enhancement	189
6.3.7	Use of Principal Components for Image Compression	193
6.3.8	The Principal Components Transform in Change Detection Applications	194
6.3.9	Use of Principal Components for Feature Reduction	198
6.4	The Noise Adjusted Principal Components Transform	198
6.5	The Kauth-Thomas Tasseled Cap Transform	201
6.6	The Kernel Principal Components Transform	205
6.7	HSI Image Display	208
6.8	Pan Sharpening	210
6.9	Bibliography on Spectral Domain Image Transforms	211
6.10	Problems	212
7	Spatial Domain Image Transforms	217
7.1	Introduction	217
7.2	Special Functions	218
7.2.1	The Complex Exponential Function	218
7.2.2	The Impulse or Delta Function	220
7.2.3	The Heaviside Step Function	221
7.3	The Fourier Series	222
7.4	The Fourier Transform	224
7.5	The Discrete Fourier Transform	227
7.5.1	Properties of the Discrete Fourier Transform	229
7.5.2	Computing the Discrete Fourier Transform	230
7.6	Convolution	230
7.6.1	The Convolution Integral	230
7.6.2	Convolution with an Impulse	231
7.6.3	The Convolution Theorem	231
7.6.4	Discrete Convolution	232
7.7	Sampling Theory	233
7.8	The Discrete Fourier Transform of an Image	236
7.8.1	The Transform Equations	236
7.8.2	Evaluating the Fourier Transform of an Image	237
7.8.3	The Concept of Spatial Frequency	238
7.8.4	Displaying the DFT of an Image	238
7.9	Image Processing Using the Fourier Transform	239
7.10	Convolution in Two Dimensions	241
7.11	Other Fourier Transforms	242
7.12	Leakage and Window Functions	243
7.13	The Wavelet Transform	244
7.13.1	Background	244

7.13.2	Orthogonal Functions and Inner Products	245
7.13.3	Wavelets as Basis Functions	246
7.13.4	Dyadic Wavelets with Compact Support	247
7.13.5	Choosing the Wavelets	248
7.13.6	Filter Banks	248
7.13.6.1	Sub Band Filtering, and Downsampling	248
7.13.6.2	Reconstruction from the Wavelets, and Upsampling	252
7.13.6.3	Relationship Between the Low and High Pass Filters	253
7.13.7	Choice of Wavelets	254
7.14	The Wavelet Transform of an Image	256
7.15	Applications of the Wavelet Transform in Remote Sensing Image Analysis	257
7.16	Bibliography on Spatial Domain Image Transforms	259
7.17	Problems	260
8	Supervised Classification Techniques	263
8.1	Introduction	263
8.2	The Essential Steps in Supervised Classification	264
8.3	Maximum Likelihood Classification	267
8.3.1	Bayes' Classification	267
8.3.2	The Maximum Likelihood Decision Rule	267
8.3.3	Multivariate Normal Class Models	269
8.3.4	Decision Surfaces	270
8.3.5	Thresholds	271
8.3.6	Number of Training Pixels Required	273
8.3.7	The Hughes Phenomenon and the Curse of Dimensionality	274
8.3.8	An Example	276
8.4	Gaussian Mixture Models	278
8.5	Minimum Distance Classification	284
8.5.1	The Case of Limited Training Data	284
8.5.2	The Discriminant Function	285
8.5.3	Decision Surfaces for the Minimum Distance Classifier	286
8.5.4	Thresholds	286
8.5.5	Degeneration of Maximum Likelihood to Minimum Distance Classification	286
8.5.6	Classification Time Comparison of the Maximum Likelihood and Minimum Distance Rules	287
8.6	Parallelepiped Classification	288
8.7	Mahalanobis Classification	289

8.8	Non-parametric Classification	290
8.9	Table Look Up Classification	291
8.10	k NN (Nearest Neighbour) Classification	291
8.11	The Spectral Angle Mapper	292
8.12	Non-parametric Classification from a Geometric Basis	293
8.12.1	Linear Classification and the Concept of a Weight Vector	293
8.12.2	Testing Class Membership	294
8.13	Training a Linear Classifier	295
8.14	The Support Vector Machine: Linearly Separable Classes	295
8.15	The Support Vector Machine: Overlapping Classes	300
8.16	The Support Vector Machine: Nonlinearly Separable Data and Kernels	303
8.17	Multi-category Classification with Binary Classifiers	306
8.18	Applying the Support Vector Classifier	307
8.18.1	Initial Choices	307
8.18.2	Grid Searching for Parameter Determination	308
8.18.3	Data Centering and Scaling	309
8.18.4	Examples	309
8.19	Committees of Classifiers	312
8.19.1	Bagging	313
8.19.2	Boosting and AdaBoost	313
8.20	Networks of Classifiers: The Artificial Neural Network	315
8.20.1	The Processing Element	316
8.20.2	Training the Neural Network—Backpropagation	317
8.20.3	Choosing the Network Parameters	323
8.20.4	Example	323
8.21	The Convolutional Neural Network	326
8.21.1	The Basic Topology of the Convolutional Neural Network	328
8.21.2	Detecting Spatial Structure	332
8.21.3	Stride	332
8.21.4	Pooling or Down-Sampling	333
8.21.5	The ReLU Activation Function	333
8.21.6	Handling the Outputs of a CNN	334
8.21.7	Multiple Filters in the Convolution Layer	335
8.21.8	Simplified Representation of the CNN	336
8.21.9	Multispectral and Hyperspectral Inputs to a CNN	336
8.21.10	A Spectral-Spatial Example of the Use of the CNN	339
8.21.11	Avoiding Overfitting	340
8.21.12	Variations	341
8.22	Recurrent Neural Networks	343

8.22.1	Multi-temporal Remote Sensing	343
8.22.2	Importance of Memory	343
8.22.3	The Recurrent Neural Network (RNN) Architecture	344
8.22.4	Training the RNN	346
8.23	Context Classification	346
8.23.1	The Concept of Spatial Context	346
8.23.2	Context Classification by Image Pre-processing	348
8.23.3	Post Classification Filtering	349
8.23.4	Probabilistic Relaxation Labelling	349
8.23.4.1	The Algorithm	349
8.23.4.2	The Neighbourhood Function	351
8.23.4.3	Determining the Compatibility Coefficients	352
8.23.4.4	Stopping the Process	352
8.23.4.5	Examples	353
8.23.5	Handling Spatial Context by Markov Random Fields	355
8.24	Bibliography on Supervised Classification Techniques	359
8.25	Problems	363
9	Clustering and Unsupervised Classification	369
9.1	How Clustering is Used	369
9.2	Similarity Metrics and Clustering Criteria	370
9.3	k Means Clustering	372
9.3.1	The k Means Algorithm	373
9.4	Isodata Clustering	374
9.4.1	Merging and Deleting Clusters	375
9.4.2	Splitting Elongated Clusters	375
9.5	Choosing the Initial Cluster Centres	375
9.6	Cost of k Means and Isodata Clustering	376
9.7	Unsupervised Classification	376
9.8	An Example of Clustering with the k Means Algorithm	377
9.9	A Single Pass Clustering Technique	378
9.9.1	The Single Pass Algorithm	379
9.9.2	Advantages and Limitations of the Single Pass Algorithm	380
9.9.3	Strip Generation Parameter	381
9.9.4	Variations on the Single Pass Algorithm	381
9.9.5	An Example of Clustering with the Single Pass Algorithm	381
9.10	Hierarchical Clustering	383
9.10.1	Agglomerative Hierarchical Clustering	383
9.11	Other Clustering Metrics	383
9.12	Some Alternative Clustering Techniques	385

9.12.1	Histogram Peak Selection	385
9.12.2	Mountain Clustering	385
9.12.3	k Medians Clustering	386
9.12.4	k Medoids Clustering	386
9.13	Clustering Large Data Sets	388
9.13.1	The K Trees Algorithm	389
9.13.2	DBSCAN	393
9.14	Cluster Space Classification	395
9.15	Bibliography on Clustering and Unsupervised Classification	399
9.16	Problems	400
10	Feature Reduction	403
10.1	The Need for Feature Reduction	403
10.2	Approaches to Feature Reduction	405
10.3	Feature Reduction by Spectral Transforms	406
10.3.1	Feature Reduction Using the Principal Components Transform	406
10.3.2	Feature Reduction Using the Canonical Analysis Transform	409
10.3.2.1	Within-Class and Among-Class Covariance	409
10.3.2.2	A Separability Measure	411
10.3.2.3	The Generalised Eigenvalue Equation	411
10.3.2.4	An Example	413
10.3.3	Discriminant Analysis Feature Extraction (DAFE)	415
10.3.4	Non-parametric Discriminant Analysis (NDA)	417
10.3.5	Decision Boundary Feature Extraction (DBFE)	421
10.3.6	Non-parametric Weighted Feature Extraction (NWFE)	422
10.4	Feature Reduction by Block Diagonalising the Covariance Matrix	423
10.5	Feature Selection	429
10.5.1	Measures of Separability	429
10.5.2	Divergence	430
10.5.2.1	Definition	430
10.5.2.2	Divergence of a Pair of Normal Distributions	432
10.5.2.3	Using Divergence for Feature Selection	432
10.5.2.4	A Problem with Divergence	433
10.5.3	The Jeffries-Matusita (JM) Distance	433
10.5.3.1	Definition	433

	10.5.3.2	Comparison of Divergence and JM Distance	435
10.5.4		Transformed Divergence	436
	10.5.4.1	Definition	436
	10.5.4.2	Transformed Divergence and the Probability of Correct Classification	436
	10.5.4.3	Use of Transformed Divergence in Clustering	437
10.5.5		Separability Measures for Minimum Distance Classification	437
10.6		Distribution Free Feature Selection—ReliefF	438
10.7		Improving Covariance Estimates Through Regularisation	440
10.8		Bibliography on Feature Reduction	442
10.9		Problems	443
11		Image Classification in Practice	447
11.1		Introduction	447
11.2		An Overview of Classification	448
	11.2.1	Supervised Classification	448
	11.2.1.1	Selection of Training Data	448
	11.2.1.2	Feature Selection	449
	11.2.1.3	Classifier Outputs and Accuracy Checking	450
	11.2.2	Unsupervised Classification	450
	11.2.3	Semi-supervised Classification and Transfer Learning	452
11.3		Effect of Resampling on Classification	453
11.4		A Hybrid Supervised/Unsupervised Methodology	454
	11.4.1	Outline of the Method	454
	11.4.2	Choosing the Image Segments to Cluster	455
	11.4.3	Rationalising the Number of Spectral Classes	456
	11.4.4	An Example	456
	11.4.5	Hybrid Classification with Other Supervised Algorithms	459
11.5		Cluster Space Classification	461
11.6		Assessing Classification Accuracy	462
	11.6.1	Use of a Testing Set of Pixels	462
	11.6.2	The Error Matrix	463
	11.6.3	Quantifying the Error Matrix	464
	11.6.4	The Kappa Coefficient	467
	11.6.5	Number of Testing Samples Required for Assessing Map Accuracy	472
	11.6.6	Number of Testing Samples Required for Populating the Error Matrix	476

11.6.7	Placing Confidence Limits on Assessed Accuracy	478
11.6.8	Cross Validation Accuracy Assessment and the Leave One Out Method	479
11.7	Decision Tree Classifiers	479
11.7.1	CART (Classification and Regression Trees)	482
11.7.2	Random Forests	485
11.7.3	Progressive Two-Class Decision Classifier	487
11.8	Image Interpretation Through Spectroscopy and Spectral Library Searching	488
11.9	End Members and Unmixing	490
11.10	Is There a Best Classifier?	492
11.10.1	Segmenting the Spectral Space	492
11.10.2	Comparing the Classifiers	494
11.11	Bibliography on Image Classification in Practice	497
11.12	Problems	500
12	Multisource Image Analysis	503
12.1	Introduction	503
12.2	Stacked Vector Analysis	504
12.3	Statistical Multisource Methods	505
12.3.1	Joint Statistical Decision Rules	505
12.3.2	Committee Classifiers	507
12.3.3	Opinion Pools and Consensus Theory	508
12.3.4	Use of Prior Probabilities	509
12.3.5	Supervised Label Relaxation	510
12.4	The Theory of Evidence	510
12.4.1	The Concept of Evidential Mass	511
12.4.2	Combining Evidence with the Orthogonal Sum	513
12.4.3	Decision Rules	515
12.5	Knowledge-Based Image Analysis	515
12.5.1	Emulating Photointerpretation to Understand Knowledge Processing	516
12.5.2	The Structure of a Knowledge-Based Image Analysis System	517
12.5.3	Representing Knowledge in a Knowledge-Based Image Analysis System	518
12.5.4	Processing Knowledge—The Inference Engine	520
12.5.5	Rules as Justifiers of a Labelling Proposition	521
12.5.6	Endorsing a Labelling Proposition	522
12.5.7	An Example	523
12.6	Operational Multisource Analysis	525
12.7	Bibliography on Multisource Image Analysis	528
12.8	Problems	530

Appendices	533
Appendix A: Satellite Altitudes and Periods	535
Appendix B: Binary Representation of Decimal Numbers	537
Appendix C: Essential Results from Vector and Matrix Algebra	539
Appendix D: Some Fundamental Material from Probability and Statistics	551
Appendix E: Penalty Function Derivation of the Maximum Likelihood Decision Rule	555
Index	559