CONTENTS

I IMAGE FORMATION **1**

1 RADIOMETRY — MEASURING LIGHT **3**

1.1 Light in Space 3

1.1.1 Foreshortening 3

1.1.2 Solid Angle 4

1.1.3 Radiance 6

1.2 Light at Surfaces 8

1.2.1 Simplifying Assumptions 9

1.2.2 The Bidirectional Reflectance Distribution Function 9

1.3 Important Special Cases 11

1.3.1 Radiosity 11

1.3.2 Directional Hemispheric Reflectance 12

1.3.3 Lambertian Surfaces and Albedo 12

1.3.4 Specular Surfaces 13

1.3.5 The Lambertian + Specular Model 14

1.4 Quick Reference: Radiometric Terminology for Light 16

1.5 Quick Reference: Radiometric Properties of Surfaces 17

1.6 Quick Reference: Important Types of Surface 18

1.7 Notes 19

1.8 Assignments 19

2 SOURCES, SHADOWS AND SHADING **21**

2.1 Radiometric Properties of Light Sources 21

2.2 Qualitative Radiometry 22

2.3 Sources and their Effects 23

2.3.1 Point Sources 24

2.3.2 Line Sources 26
2.3.3 Area Sources 27
2.4 Local Shading Models 28
2.4.1 Local Shading Models for Point Sources 28
2.4.2 Area Sources and their Shadows 31
2.4.3 Ambient Illumination 31
2.5 Application: Photometric Stereo 33
2.5.1 Normal and Albedo from Many Views 36
2.5.2 Shape from Normals 37
2.6 Interreflections: Global Shading Models 40
2.6.1 An Interreflection Model 42
2.6.2 Solving for Radiosity 43
2.6.3 The qualitative effects of interreflections 45
2.7 Notes 47
2.8 Assignments 50
2.8.1 Exercises 50
2.8.2 Programming Assignments 51

3 COLOUR 53
3.1 The Physics of Colour 53
3.1.1 Radiometry for Coloured Lights: Spectral Quantities 53
3.1.2 The Colour of Surfaces 54
3.1.3 The Colour of Sources 55
3.2 Human Colour Perception 58
3.2.1 Colour Matching 58
3.2.2 Colour Receptors 61
3.3 Representing Colour 63
3.3.1 Linear Colour Spaces 63
3.3.2 Non-linear Colour Spaces 68
3.3.3 Spatial and Temporal Effects 73
3.4 Application: Finding Specularities 73
3.5 Surface Colour from Image Colour 77
3.5.1 Surface Colour Perception in People 77
3.5.2 Inferring Lightness 80
3.5.3 A Model for Image Colour 83
3.5.4 Surface Colour from Finite Dimensional Linear Models 86
3.6 Notes 89

3.6.1 Trichromacy and Colour Spaces 89
3.6.2 Lightness and Colour Constancy 90
3.6.3 Colour in Recognition 91
3.7 Assignments 91

II IMAGE MODELS 94

4 GEOMETRIC IMAGE FEATURES 96
4.1 Elements of Differential Geometry 100
4.1.1 Curves 100
4.1.2 Surfaces 105
Application: The shape of specularities 109
4.2 Contour Geometry 112
4.2.1 The Occluding Contour and the Image Contour 113
4.2.2 The Cusps and Inflections of the Image Contour 114
4.2.3 Koenderink's Theorem 115
4.3 Notes 117
4.4 Assignments 118

5 ANALYTICAL IMAGE FEATURES 120
5.1 Elements of Analytical Euclidean Geometry 120
5.1.1 Coordinate Systems and Homogeneous Coordinates 121
5.1.2 Coordinate System Changes and Rigid Transformations 124
5.2 Geometric Camera Parameters 129
5.2.1 Intrinsic Parameters 129
5.2.2 Extrinsic Parameters 132
5.2.3 A Characterization of Perspective Projection Matrices 132
5.3 Calibration Methods 133
5.3.1 A Linear Approach to Camera Calibration 134
Technique: Linear Least Squares Methods 135
5.3.2 Taking Radial Distortion into Account 139
5.3.3 Using Straight Lines for Calibration 140
5.3.4 Analytical Photogrammetry 143
Technique: Non-Linear Least Squares Methods 145
5.4 Notes 147
5.5 Assignments 147

6 AN INTRODUCTION TO PROBABILITY 150
6.1 Probability in Discrete Spaces 151
6.1.1 Probability: the P-function 151
6.1.2 Conditional Probability 153
6.1.3 Choosing P 153
6.2 Probability in Continuous Spaces 159
6.2.1 Event Structures for Continuous Spaces 159
6.2.2 Representing a P-function for the Real Line 160
6.2.3 Probability Densities 161
6.3 Random Variables 161
6.3.1 Conditional Probability and Independence 162
6.3.2 Expectations 163
6.3.3 Joint Distributions and Marginalization 165
6.4 Standard Distributions and Densities 165
6.4.1 The Normal Distribution 167
6.5 Probabilistic Inference 167
6.5.1 The Maximum Likelihood Principle 168
6.5.2 Priors, Posteriors and Bayes' rule 170
6.5.3 Bayesian Inference 170
6.5.4 Open Issues 177
6.6 Discussion 178

III EARLY VISION: ONE IMAGE 180

7 LINEAR FILTERS 182
7.1 Linear Filters and Convolution 182
7.1.1 Convolution 182
7.1.2 Example: Smoothing by Averaging 183
7.1.3 Example: Smoothing with a Gaussian 185
7.2 Shift invariant linear systems 186
7.2.1 Discrete Convolution 188
7.2.2 Continuous Convolution 190
7.2.3 Edge Effects in Discrete Convolutions 192
7.3 Spatial Frequency and Fourier Transforms 193
7.3.1 Fourier Transforms 193
7.4 Sampling and Aliasing 197
7.4.1 Sampling 198

7.4.2 Aliasing 201
7.4.3 Smoothing and Resampling 202
7.5 Technique: Scale and Image Pyramids 204
7.5.1 The Gaussian Pyramid 205
7.5.2 Applications of Scaled Representations 206
7.5.3 Scale Space 208
7.6 Discussion 211
7.6.1 Real Imaging Systems vs Shift-Invariant Linear Systems 211
7.6.2 Scale 212

8 EDGE DETECTION 214
8.1 Estimating Derivatives with Finite Differences 214
8.1.1 Differentiation and Noise 216
8.1.2 Laplacians and edges 217
8.2 Noise 217
8.2.1 Additive Stationary Gaussian Noise 219
8.3 Edges and Gradient-based Edge Detectors 224
8.3.1 Estimating Gradients 224
8.3.2 Choosing a Smoothing Filter 225
8.3.3 Why Smooth with a Gaussian? 227
8.3.4 Derivative of Gaussian Filters 229
8.3.5 Identifying Edge Points from Filter Outputs 230
8.4 Commentary 234

9 FILTERS AND FEATURES 237
9.1 Filters as Templates 237
9.1.1 Convolution as a Dot Product 237
9.1.2 Changing Basis 238
9.2 Human Vision: Filters and Primate Early Vision 239
9.2.1 The Visual Pathway 239
9.2.2 How the Visual Pathway is Studied 241
9.2.3 The Response of Retinal Cells 241
9.2.4 The Lateral Geniculate Nucleus 242
9.2.5 The Visual Cortex 243
9.2.6 A Model of Early Spatial Vision 246
9.3 Technique: Normalised Correlation and Finding Patterns 248
9.3.1 Controlling the Television by Finding Hands by Normalised Correlation 248

9.4 Corners and Orientation Representations 249
9.5 Advanced Smoothing Strategies and Non-linear Filters 252
9.5.1 More Noise Models 252
9.5.2 Robust Estimates 253
9.5.3 Median Filters 254
9.5.4 Mathematical morphology: erosion and dilation 257
9.5.5 Anisotropic Scaling 258
9.6 Commentary 259

10 TEXTURE 261
10.1 Representing Texture 263
10.1.1 Extracting Image Structure with Filter Banks 263
10.2 Analysis (and Synthesis) Using Oriented Pyramids 268
10.2.1 The Laplacian Pyramid 269
10.2.2 Oriented Pyramids 272
10.3 Application: Synthesizing Textures for Rendering 272
10.3.1 Homogeneity 274
10.3.2 Synthesis by Matching Histograms of Filter Responses 275
10.3.3 Synthesis by Sampling Conditional Densities of Filter Responses 280
10.3.4 Synthesis by Sampling Local Models 284
10.4 Shape from Texture: Planes and Isotropy 286
10.4.1 Recovering the Orientation of a Plane from an Isotropic Texture 288
10.4.2 Recovering the Orientation of a Plane from an Homogeneity Assumption 290
10.4.3 Shape from Texture for Curved Surfaces 291
10.5 Notes 292
10.5.1 Shape from Texture 293

IV EARLY VISION: MULTIPLE IMAGES 295

11 THE GEOMETRY OF MULTIPLE VIEWS 297
11.1 Two Views 298
11.1.1 Epipolar Geometry 298
11.1.2 The Calibrated Case 299
11.1.3 Small Motions 300
11.1.4 The Uncalibrated Case 301
11.1.5 Weak Calibration 302

11.2 Three Views 305
11.2.1 Trifocal Geometry 307
11.2.2 The Calibrated Case 307
11.2.3 The Uncalibrated Case 309
11.2.4 Estimation of the Trifocal Tensor 310
11.3 More Views 311
11.4 Notes 317
11.5 Assignments 319

12 STEREOPSIS 321
12.1 Reconstruction 323
12.1.1 Camera Calibration 324
12.1.2 Image Rectification 325
Human Vision: Stereopsis 327
12.2 Binocular Fusion 331
12.2.1 Correlation 331
12.2.2 Multi-Scale Edge Matching 333
12.2.3 Dynamic Programming 336
12.3 Using More Cameras 338
12.3.1 Trinocular Stereo 338
12.3.2 Multiple-Baseline Stereo 340
12.4 Notes 341
12.5 Assignments 343

13 AFFINE STRUCTURE FROM MOTION 345
13.1 Elements of Affine Geometry 346
13.2 Affine Structure from Two Images 349
13.2.1 The Affine Structure-from-Motion Theorem 350
13.2.2 Rigidity and Metric Constraints 351
13.3 Affine Structure from Multiple Images 351
13.3.1 The Affine Structure of Affine Image Sequences 352
Technique: Singular Value Decomposition 353
13.3.2 A Factorization Approach to Affine Motion Analysis 353
13.4 From Affine to Euclidean Images 356
13.4.1 Euclidean Projection Models 357
13.4.2 From Affine to Euclidean Motion 358
13.5 Affine Motion Segmentation 360
13.5.1 The Reduced Echelon Form of the Data Matrix 360

13.5.2 The Shape Interaction Matrix 360
13.6 Notes 362
13.7 Assignments 363

14 PROJECTIVE STRUCTURE FROM MOTION 365
14.1 Elements of Projective Geometry 366
14.1.1 Projective Bases and Projective Coordinates 366
14.1.2 Projective Transformations 368
14.1.3 Affine and Projective Spaces 370
14.1.4 Hyperplanes and Duality 371
14.1.5 Cross-Ratios 372
14.1.6 Application: Parameterizing the Fundamental Matrix 375
14.2 Projective Scene Reconstruction from Two Views 376
14.2.1 Analytical Scene Reconstruction 376
14.2.2 Geometric Scene Reconstruction 378
14.3 Motion Estimation from Two or Three Views 379
14.3.1 Motion Estimation from Fundamental Matrices 379
14.3.2 Motion Estimation from Trifocal Tensors 381
14.4 Motion Estimation from Multiple Views 382
14.4.1 A Factorization Approach to Projective Motion Analysis 383
14.4.2 Bundle Adjustment 386
14.5 From Projective to Euclidean Structure and Motion 386
14.5.1 Metric Upgrades from (Partial) Camera Calibration 387
14.5.2 Metric Upgrades from Minimal Assumptions 389
14.6 Notes 392
14.7 Assignments 394

V MID-LEVEL VISION 399

15 SEGMENTATION USING CLUSTERING METHODS 401
15.1 Human vision: Grouping and Gestalt 403
15.2 Applications: Shot Boundary Detection, Background Subtraction and Skin Finding 407
15.2.1 Background Subtraction 407
15.2.2 Shot Boundary Detection 408
15.2.3 Finding Skin Using Image Colour 410
15.3 Image Segmentation by Clustering Pixels 411

15.3.1 Simple Clustering Methods 411
15.3.2 Segmentation Using Simple Clustering Methods 413
15.3.3 Clustering and Segmentation by K-means 415
15.4 Segmentation by Graph-Theoretic Clustering 417
15.4.1 Basic Graphs 418
15.4.2 The Overall Approach 420
15.4.3 Affinity Measures 420
15.4.4 Eigenvectors and Segmentation 424
15.4.5 Normalised Cuts 427
15.5 Discussion 430

16 FITTING 436
16.1 The Hough Transform 437
16.1.1 Fitting Lines with the Hough Transform 437
16.1.2 Practical Problems with the Hough Transform 438
16.2 Fitting Lines 440
16.2.1 Least Squares, Maximum Likelihood and Parameter Estimation 441
16.2.2 Which Point is on Which Line? 444
16.3 Fitting Curves 445
16.3.1 Implicit Curves 446
16.3.2 Parametric Curves 449
16.4 Fitting to the Outlines of Surfaces 450
16.4.1 Some Relations Between Surfaces and Outlines 451
16.4.2 Clustering to Form Symmetries 453
16.5 Discussion 457

17 SEGMENTATION AND FITTING USING PROBABILISTIC METHODS 460
17.1 Missing Data Problems, Fitting and Segmentation 461
17.1.1 Missing Data Problems 461
17.1.2 The EM Algorithm 463
17.1.3 Colour and Texture Segmentation with EM 469
17.1.4 Motion Segmentation and EM 470
17.1.5 The Number of Components 474
17.1.6 How Many Lines are There? 474
17.2 Robustness 475
17.2.1 Explicit Outliers 475
17.2.2 M-estimators 477

17.2.3 RANSAC 480
17.3 How Many are There? 483
17.3.1 Basic Ideas 484
17.3.2 AIC — An Information Criterion 484
17.3.3 Bayesian methods and Schwartz' BIC 485
17.3.4 Description Length 486
17.3.5 Other Methods for Estimating Deviance 486
17.4 Discussion 487

18 TRACKING 489
18.1 Tracking as an Abstract Inference Problem 490
18.1.1 Independence Assumptions 490
18.1.2 Tracking as Inference 491
18.1.3 Overview 492
18.2 Linear Dynamic Models and the Kalman Filter 492
18.2.1 Linear Dynamic Models 492
18.2.2 Kalman Filtering 497
18.2.3 The Kalman Filter for a 1D State Vector 497
18.2.4 The Kalman Update Equations for a General State Vector 499
18.2.5 Forward-Backward Smoothing 500
18.3 Non-Linear Dynamic Models 505
18.3.1 Unpleasant Properties of Non-Linear Dynamics 508
18.3.2 Difficulties with Likelihoods 509
18.4 Particle Filtering 511
18.4.1 Sampled Representations of Probability Distributions 511
18.4.2 The Simplest Particle Filter 515
18.4.3 A Workable Particle Filter 518
18.4.4 If's, And's and But's — Practical Issues in Building Particle Filters 519
18.5 Data Association 523
18.5.1 Choosing the Nearest — Global Nearest Neighbours 523
18.5.2 Gating and Probabilistic Data Association 524
18.6 Applications and Examples 527
18.6.1 Vehicle Tracking 528
18.6.2 Finding and Tracking People 532
18.7 Discussion 538
II Appendix: The Extended Kalman Filter, or EKF 540

VI HIGH-LEVEL VISION 542

19 CORRESPONDENCE AND POSE CONSISTENCY 544
19.1 Initial Assumptions 544
19.1.1 Obtaining Hypotheses 545
19.2 Obtaining Hypotheses by Pose Consistency 546
19.2.1 Pose Consistency for Perspective Cameras 547
19.2.2 Affine and Projective Camera Models 549
19.2.3 Linear Combinations of Models 551
19.3 Obtaining Hypotheses by Pose Clustering 553
19.4 Obtaining Hypotheses Using Invariants 554
19.4.1 Invariants for Plane Figures 554
19.4.2 Geometric Hashing 559
19.4.3 Invariants and Indexing 560
19.5 Verification 564
19.5.1 Edge Proximity 565
19.5.2 Similarity in Texture, Pattern and Intensity 567
19.5.3 Example: Bayes Factors and Verification 567
19.6 Application: Registration in Medical Imaging Systems 568
19.6.1 Imaging Modes 569
19.6.2 Applications of Registration 570
19.6.3 Geometric Hashing Techniques in Medical Imaging 571
19.7 Curved Surfaces and Alignment 573
19.8 Discussion 576

20 FINDING TEMPLATES USING CLASSIFIERS 581
20.1 Classifiers 582
20.1.1 Using Loss to Determine Decisions 582
20.1.2 Overview: Methods for Building Classifiers 584
20.1.3 Example: A Plug-in Classifier for Normal Class-conditional Densities 586
20.1.4 Example: A Non-Parametric Classifier using Nearest Neighbours 587
20.1.5 Estimating and Improving Performance 588
20.2 Building Classifiers from Class Histograms 590
20.2.1 Finding Skin Pixels using a Classifier 591
20.2.2 Face Finding Assuming Independent Template Responses 592
20.3 Feature Selection 595

20.3.1 Principal Component Analysis 595
20.3.2 Canonical Variates 597
20.4 Neural Networks 601
20.4.1 Key Ideas 601
20.4.2 Minimizing the Error 606
20.4.3 When to Stop Training 610
20.4.4 Finding Faces using Neural Networks 610
20.4.5 Convolutional Neural Nets 612
20.5 The Support Vector Machine 615
20.5.1 Support Vector Machines for Linearly Separable Datasets 616
20.5.2 Finding Pedestrians using Support Vector Machines 618
20.6 Conclusions 622
II Appendix: Support Vector Machines for Datasets that are not Linearly Separable 624
III Appendix: Using Support Vector Machines with Non-Linear Kernels 625

21 RECOGNITION BY RELATIONS BETWEEN TEMPLATES 627
21.1 Finding Objects by Voting on Relations between Templates 628
21.1.1 Describing Image Patches 628
21.1.2 Voting and a Simple Generative Model 629
21.1.3 Probabilistic Models for Voting 630
21.1.4 Voting on Relations 632
21.1.5 Voting and 3D Objects 632
21.2 Relational Reasoning using Probabilistic Models and Search 633
21.2.1 Correspondence and Search 633
21.2.2 Example: Finding Faces 636
21.3 Using Classifiers to Prune Search 639
21.3.1 Identifying Acceptable Assemblies Using Projected Classifiers 640
21.3.2 Example: Finding People and Horses Using Spatial Relations 640
21.4 Technique: Hidden Markov Models 643
21.4.1 Formal Matters 644
21.4.2 Computing with Hidden Markov Models 645
21.4.3 Varieties of HMM's 652
21.5 Application: Hidden Markov Models and Sign Language Understanding 654
21.6 Application: Finding People with Hidden Markov Models 659
21.7 Frames and Probability Models 662
21.7.1 Representing Coordinate Frames Explicitly in a Probability Model 664

21.7.2 Using a Probability Model to Predict Feature Positions 666
21.7.3 Building Probability Models that are Frame-Invariant 668
21.7.4 Example: Finding Faces Using Frame Invariance 669
21.8 Conclusions 669

22 ASPECT GRAPHS 672
22.1 Differential Geometry and Visual Events 677
22.1.1 The Geometry of the Gauss Map 677
22.1.2 Asymptotic Curves 679
22.1.3 The Asymptotic Spherical Map 681
22.1.4 Local Visual Events 682
22.1.5 The Bitangent Ray Manifold 684
22.1.6 Multilocal Visual Events 686
22.1.7 Remarks 687
22.2 Computing the Aspect Graph 689
22.2.1 Step 1: Tracing Visual Events 690
22.2.2 Step 2: Constructing the Regions 691
22.2.3 Remaining Steps of the Algorithm 692
22.2.4 An Example 692
22.3 Aspect Graphs and Object Recognition 696
22.4 Notes 696
22.5 Assignments 697

VII APPLICATIONS AND TOPICS 699

23 RANGE DATA 701
23.1 Active Range Sensors 701
23.2 Range Data Segmentation 704
Technique: Analytical Differential Geometry 705
23.2.1 Finding Step and Roof Edges in Range Images 707
23.2.2 Segmenting Range Images into Planar Regions 712
23.3 Range Image Registration and Model Construction 714
Technique: Quaternions 715
23.3.1 Registering Range Images Using the Iterative Closest-Point Method 716
23.3.2 Fusing Multiple Range Images 719
23.4 Object Recognition 720

23.4.1 Matching Piecewise-Planar Surfaces Using Interpretation Trees 721
23.4.2 Matching Free-Form Surfaces Using Spin Images 724
23.5 Notes 729
23.6 Assignments 730

24 APPLICATION: FINDING IN DIGITAL LIBRARIES 732
24.1 Background 733
24.1.1 What do users want? 733
24.1.2 What can tools do? 735
24.2 Appearance 736
24.2.1 Histograms and correlograms 737
24.2.2 Textures and textures of textures 738
24.3 Finding 745
24.3.1 Annotation and segmentation 748
24.3.2 Template matching 749
24.3.3 Shape and correspondence 751
24.4 Video 754
24.5 Discussion 756

25 APPLICATION: IMAGE-BASED RENDERING 758
25.1 Constructing 3D Models from Image Sequences 759
25.1.1 Scene Modeling from Registered Images 759
25.1.2 Scene Modeling from Unregistered Images 767
25.2 Transfer-Based Approaches to Image-Based Rendering 771
25.2.1 Affine View Synthesis 772
25.2.2 Euclidean View Synthesis 775
25.3 The Light Field 778
25.4 Notes 782
25.5 Assignments 784