

Course Syllabus

ECE 6930 - Special Topics in Electrical Engineering

“Machine Vision in Control and Automation”

Summer 2002, Independent Study Course.

Instructor: YangQuan Chen, Center for Self-Organizing and Intelligent Systems
Department of Electrical and Computer Engineering, Utah State University
Room EL-256. T: (435) 797-0148, E: yqchen@ece.usu.edu

Consulting Professor: Kevin L. Moore.

Office Hours: Friday 3:00 PM to 5:30 PM or by appointment.

Text: D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Prentice Hall, 2003. (Early version of the e-copy is available, 973 pages.)

References:

Books:

- [1]. L. G. Shapiro and G. C. Stockman, *Computer Vision*, Prentice Hall, 2001.
- [2]. A. Watt, *The Computer Image*, Addison-Wesley, 1998.
- [3]. M. Sonka, V. Hlavac and R. Boyle. *Image Processing, Analysis, and Machine Vision*, Brooks/Cole Publishing, 1999.
- [4]. E. Trucco and A. Verri. *Introductory Techniques for 3-D Computer Vision*, Prentice-Hall, 1998.
- [5]. R. Jain, R. Kasturi, and B. G. Schunck. *Machine Vision*, MacGraw-Hill, 1995.
- [6]. V. S. Nalwa. *A Guided Tour of Computer Vision*, Addison-Wesley, 1993.
- [7]. O. Faugeras. *Three-Dimensional Computer Vision*, MIT Press, 1993.

Survival URLs:

- <http://www-2.cs.cmu.edu/afs/cs/project/cil/ftp/html/vision.html>
- <http://www.dai.ed.ac.uk/CVonline/>
- <http://iris.usc.edu/Vision-Notes/bibliography/contents.html>

Journals

1. Computer vision and image understanding : CVIU
(USU SCITECH Call #: TA 1632 .C835)
2. Graphical models and image processing : GMIP
(USU SCITECH Call #: T 385 .C89)
3. CVGIP. Graphical models and image processing
(USU SCITECH Call #: T 385 .C83)
4. International journal of computer vision
(USU SCITECH Call #: TA 1632 .I57)
5. IEEE transactions on pattern analysis and machine intelligence (PAMI)
(USU SCITECH Call #: Q 334 .I43)

Prerequisites: Undergraduate control systems (ECE4310, ECE5320) and graduate linear multivariable systems (ECE6320). A good working knowledge of C and C++ programming and Matlab/Simulink, Linear algebra, Vector calculus etc. No prior knowledge of vision is assumed. Knowledge with “Nonlinear and adaptive control” (ECE6330) will be an added advantage but not compulsory.

Credits: 3

Course Load: 10 hours per week. 12 weeks.

Course Requirements:

4 Projects	80 points
2 Literature Survey Reports:	20 points

(The details will be sent via email.)
There is no Mid-term Exam and Final Exam.

Notes:

1. This course is designed for 12 weeks summer independent study.
2. To get 3 credits, you have to spend at least 10 hours a week and 120 hours in total in this course, although you have the *flexibility* to arrange your efforts.
3. A weekly group discussion is encouraged (but not compulsory) in the Instructor’s office in the course office hour. Friday 3:00pm-5:30pm.
4. Computer simulations will be necessary for some projects. Matlab/Simulink is the preferred computing environment for these simulations. In some projects, C/C++ programming is required.
5. *A CD containing the essential course materials is available from the Instructor.*

Course Description:

The goal of machine vision is to compute properties of the three-dimensional world from digital images. Problems in this field include identifying the 3D shape of an environment, determining how things are moving, and recognizing familiar people and objects, all through analysis of images and video. However, in control and automation applications, we are more interested in using vision to achieve better control performance for certain tasks not possible by other means.

In addition to common topics in machine vision such as feature detection, image segmentation, motion estimation and object recognition, we will emphasize on

- 1) *Visual servo problems such as visual servoing, visual docking, visual platooning, visual formation etc.*
- 2) *Template matching for applications such as visual inspection tasks in factory automation, picture retrieval in graphical database etc.*
- 3) *Camera calibration techniques for visual metrology applications.*

Appendix: Basic Syllabus and the Table of Contents

Extracted from the Text: D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Prentice Hall, 2003. (Early version of the e-copy is available, 973 pages.)

One-Semester Introductory Class

This one-semester introduction to computer vision is designed for seniors or first-year graduate students in computer science, electrical engineering, or other engineering or science disciplines. The students receive a broad presentation of the field, including application areas such as digital libraries and HCI. Although the hardest theoretical material in differential geometry and probability theory is omitted, there is a thorough treatment of the basic geometry and physics of image formation. Note that this syllabus assumes that only parts of certain chapters be covered (e.g., only the introduction and Section 1 of Chapter 1, etc.).

1. Chapter 2 Radiometry
 2. Chapter 3 Sources, Shadows and Shading
 3. Chapter 4 Color
 4. Chapter 1 Cameras
 5. Chapter 8 and Chapter 9 Linear Filters and Edge Detection
 6. Chapter 11 Texture
 7. Chapter 22 Digital Libraries
 8. Chapter 12 The Geometry of Multiple Views
 9. Chapter 13 Stereopsis
 10. Chapter 16 and Chapter 17 Segmentation and Fitting
 11. Chapter 19 Tracking Using Linear Dynamic Models
 12. Chapter 24 Correspondence and Pose
 13. Chapter 25 Template Matching
 14. Chapter 26 Recognition by Relations between Templates
 15. Chapter 29 Toward Category-Level Recognition
- and the enterprising may wish to read some of the other chapters on their own time.

Note that: the above should be considered as the “*bottom-line reading load*” for ECE6930 summer 2002 independent study, in addition to the focused reading and literature review tasks.

CONTENTS

PREFACE xxi

I IMAGE FORMATION 1

1 CAMERAS 3

1.1 Pinhole Cameras 4

1.1.1 Perspective Projection 4

1.1.2 Affine Projection 6

- 1.1.3 Spherical Projection 8
- 1.2 Cameras with Lenses 9
 - 1.2.1 First-Order Geometric Optics 11
 - 1.2.2 Thin Lenses: Geometry 12
 - 1.2.3 Thin Lenses: Radiometry 15
 - 1.2.4 Real Lenses 16
- Human Vision: The Structure of the Eye 20
- 1.3 Sensing 22
 - 1.3.1 CCD cameras 23
 - 1.3.2 Sensor Models 24
- 1.4 Notes 25
- 1.5 Assignments 27

- 2 RADIOMETRY —MEASURING LIGHT 28
 - 2.1 Light in Space 28
 - 2.1.1 Foreshortening 28
 - 2.1.2 Solid Angle 29
 - 2.1.3 Radiance 31
 - 2.2 Light at Surfaces 33
 - 2.2.1 Simplifying Assumptions 34
 - 2.2.2 The Bidirectional Reflectance Distribution Function 34
 - 2.3 Important Special Cases 36
 - 2.3.1 Radiosity 36
 - 2.3.2 Directional Hemispheric Reflectance 37
 - 2.3.3 Lambertian Surfaces and Albedo 37
 - 2.3.4 Specular Surfaces 38
 - 2.3.5 The Lambertian + Specular Model 39
 - 2.4 Quick Reference: Radiometric Terminology for Light 41
 - 2.5 Quick Reference: Radiometric Properties of Surfaces 42
 - 2.6 Quick Reference: Important Types of Surface 43
 - 2.7 Comments 44
 - 2.8 Assignments 45

- 3 SOURCES, SHADOWS AND SHADING 47
 - 3.1 Radiometric Properties of Light Sources 47
 - 3.2 Qualitative Radiometry 48
 - 3.3 Sources and their Effects 49
 - 3.3.1 Point Sources 50
 - 3.3.2 Line Sources 52
 - 3.3.3 Area Sources 53
 - 3.4 Local Shading Models 54
 - 3.4.1 Local Shading Models for Point Sources 54
 - 3.4.2 Area Sources and their Shadows 57
 - 3.4.3 Ambient Illumination 57
 - 3.5 Application: Photometric Stereo 59
 - 3.5.1 Normal and Albedo from Many Views 62
 - 3.5.2 Shape from Normals 63
 - 3.6 Interreflections: Global Shading Models 66
 - 3.6.1 An Interreflection Model 68
 - 3.6.2 Solving for Radiosity 69
 - 3.6.3 The qualitative effects of interreflections 71
 - 3.7 Notes 74
 - 3.7.1 Local Shading Models 74
 - 3.7.2 Interreflections 75
 - 3.7.3 Photometric Stereo 75
 - 3.7.4 Alternative Shading Representations 76
 - 3.8 Assignments 77
 - 3.8.1 Exercises 77
 - 3.8.2 Programming Assignments 78

- 4 COLOUR 80
 - 4.1 The Physics of Colour 80
 - 4.1.1 Radiometry for Coloured Lights: Spectral Quantities 80
 - 4.1.2 The Colour of Surfaces 81
 - 4.1.3 The Colour of Sources 82

- 4.2 Human Colour Perception 85
 - 4.2.1 Colour Matching 85
 - 4.2.2 Colour Receptors 88
- 4.3 Representing Colour 90
 - 4.3.1 Linear Colour Spaces 90
 - 4.3.2 Non-linear Colour Spaces 95
 - 4.3.3 Spatial and Temporal Effects 100
- 4.4 A Model for Image Colour 100
 - 4.4.1 Cameras 100
 - 4.4.2 A Model for Image Colour 101
 - 4.4.3 Application: Finding Specularities 105
- 4.5 Surface Colour from Image Colour 108
 - 4.5.1 Surface Colour Perception in People 109
 - 4.5.2 Inferring Lightness 110
 - 4.5.3 Surface Colour from Finite Dimensional Linear Models 115
- 4.6 Notes 118
 - 4.6.1 Trichromacy and Colour Spaces 118
 - 4.6.2 Specularity Finding 119
 - 4.6.3 Lightness 119
 - 4.6.4 Colour Constancy 120
 - 4.6.5 Colour in Recognition 121
- 5.7 Assignments 121

II IMAGE MODELS 125

5 GEOMETRIC CAMERA MODELS 127

- 5.1 Elements of Analytical Euclidean Geometry 128
 - 5.1.1 Coordinate Systems and Homogeneous Coordinates 128
 - 5.1.2 Coordinate System Changes and Rigid Transformations 132
- 5.2 Geometric Camera Parameters 137
 - 5.2.1 Intrinsic Parameters 138
 - 5.2.2 Extrinsic Parameters 140
 - 5.2.3 A Characterization of Perspective Projection Matrices 141
- 5.3 Straight Lines and their Projections 142
 - 5.3.1 Elements of Line Geometry 142
 - 5.3.2 Projection Equations 144
- 5.4 Notes 144
- 5.5 Assignments 145

6 GEOMETRIC CAMERA CALIBRATION 148

- 6.1 Least-Squares Parameter Estimation 149
 - 6.1.1 Linear Least-Squares Methods 149
 - 6.1.2 Non-Linear Least-Squares Methods 153
- 6.2 A Linear Approach to Camera Calibration 156
 - 6.2.1 Estimation of the Projection Matrix 157
 - 6.2.2 Estimation of the Intrinsic and Extrinsic Parameters 157
 - 6.2.3 Degenerate Point Configurations 158
- 6.3 Taking Radial Distortion into Account 159
 - 6.3.1 Estimation of the Projection Matrix 160
 - 6.3.2 Estimation of the Intrinsic and Extrinsic Parameters 160
 - 6.3.3 Degenerate Point Configurations 162
- 6.4 Using Straight Lines for Calibration 162
- 6.5 Analytical Photogrammetry 164
- 6.6 An Application: Mobile Robot Localization 166
- 6.7 Notes 167
- 6.8 Assignments 168

7 AN INTRODUCTION TO PROBABILITY 170

- 7.1 Probability in Discrete Spaces 171
 - 7.1.1 Probability: the P-function 172
 - 7.1.2 Conditional Probability 173
 - 7.1.3 Choosing P 174
- 7.2 Probability in Continuous Spaces 179
 - 7.2.1 Event Structures for Continuous Spaces 179
 - 7.2.2 Representing P-functions 181

- 7.2.3 Representing P-functions with Probability Density Functions 182
- 7.3 Random Variables 182
 - 7.3.1 Conditional Probability and Independence 183
 - 7.3.2 Expectations 184
 - 7.3.3 Joint Distributions and Marginalization 185
- 7.4 Standard Distributions and Densities 187
 - 7.4.1 The Normal Distribution 188
- 7.5 Probabilistic Inference 188
 - 7.5.1 The Maximum Likelihood Principle 189
 - 7.5.2 Priors, Posteriors and Bayes' rule 189
 - 7.5.3 Bayesian Inference 191
 - 7.5.4 Open Issues 198
- 7.6 Discussion 198

III EARLY VISION: ONE IMAGE 201

8 LINEAR FILTERS 203

- 8.1 Linear Filters and Convolution 203
 - 8.1.1 Convolution 204
- 8.2 Shift invariant linear systems 210
 - 8.2.1 Discrete Convolution 210
 - 8.2.2 Continuous Convolution 212
 - 8.2.3 Edge Effects in Discrete Convolutions 215
- 8.3 Spatial Frequency and Fourier Transforms 215
 - 8.3.1 Fourier Transforms 216
- 8.4 Sampling and Aliasing 219
 - 8.4.1 Sampling 220
 - 8.4.2 Aliasing 223
 - 8.4.3 Smoothing and Resampling 224
- 8.5 Technique: Scale and Image Pyramids 226
 - 8.5.1 The Gaussian Pyramid 227
 - 8.5.2 Applications of Scaled Representations 228
 - 8.5.3 Scale Space 231
- 8.6 Discussion 234
 - 8.6.1 Real Imaging Systems vs. Shift-Invariant Linear Systems 234
 - 8.6.2 Scale 235
 - 8.6.3 Anisotropic Scaling 235

9 EDGE DETECTION 238

- 9.1 Noise 238
 - 9.1.1 Additive Stationary Gaussian Noise 239
 - 9.1.2 Why Finite Differences Respond to Noise 241
- 9.2 Estimating Derivatives 243
 - 9.2.1 Choosing a Smoothing Filter 245
 - 9.2.2 Why Smooth with a Gaussian? 246
 - 9.2.3 Derivative of Gaussian Filters 249
- 9.3 Detecting Edges 249
 - 9.3.1 Using the Laplacian to Detect Edges 250
 - 9.3.2 Gradient Based Edge Detectors 251
 - 9.3.3 Technique: Orientation Representations and Corners 255
- 9.4 Commentary 260

10 FILTERS AND FEATURES 266

- 10.1 Filters as Templates 266
 - 10.1.1 Convolution as a Dot Product 266
 - 10.1.2 Changing Basis 267
- 10.2 Technique: Normalized Correlation and Finding Patterns 268
 - 10.2.1 Controlling the Television by Finding Hands by Normalized Correlation 268
- 10.3 Human Vision: Filters and Primate Early Vision 269
 - 10.3.1 The Visual Pathway 270
 - 10.3.2 The Response of Retinal Cells 272
 - 10.3.3 The Lateral Geniculate Nucleus 274
 - 10.3.4 The Visual Cortex 275
 - 10.3.5 A Model of Early Spatial Vision 278
- 10.4 Advanced Smoothing Strategies and Non-linear Filters 280

- 10.4.1 More Noise Models 280
- 10.4.2 Robust Estimates 281
- 10.4.3 Median Filters 282
- 10.4.4 Mathematical morphology: erosion and dilation 286
- 10.5 Commentary 287

11 TEXTURE 289

- 11.1 Representing Texture 290
 - 11.1.1 Extracting Image Structure with Filter Banks 291
- 11.2 Analysis (and Synthesis) Using Oriented Pyramids 294
 - 11.2.1 The Laplacian Pyramid 296
 - 11.2.2 Filters in the Spatial Frequency Domain 298
 - 11.2.3 Oriented Pyramids 302
- 11.3 Application: Synthesizing Textures for Rendering 304
 - 11.3.1 Homogeneity 306
 - 11.3.2 Synthesis by Matching Histograms of Filter Responses 306
 - 11.3.3 Synthesis by Sampling Conditional Densities of Filter Responses 309
 - 11.3.4 Synthesis by Sampling Local Models 314
- 11.4 Shape from Texture 316
 - 11.4.1 Shape from Texture for Planes 317
 - 11.4.2 Shape from Texture for Curved Surfaces 321
- 11.5 Notes 322
 - 11.5.1 Filters, Pyramids and Efficiency 323
 - 11.5.2 Texture Synthesis 323
 - 11.5.3 Shape from Texture 323

IV EARLY VISION: MULTIPLE IMAGES 326

12 THE GEOMETRY OF MULTIPLE VIEWS 328

- 12.1 Two Views 329
 - 12.1.1 Epipolar Geometry 329
 - 12.1.2 The Calibrated Case 330
 - 12.1.3 Small Motions 331
 - 12.1.4 The Uncalibrated Case 332
 - 12.1.5 Weak Calibration 333
- 12.2 Three Views 336
 - 12.2.1 Trifocal Geometry 338
 - 12.2.2 The Calibrated Case 338
 - 12.2.3 The Uncalibrated Case 340
 - 12.2.4 Estimation of the Trifocal Tensor 341
- 12.3 More Views 342
- 12.4 Notes 348
- 12.5 Assignments 350

13 STEREOPSIS 352

- 13.1 Reconstruction 354
 - 13.1.1 Camera Calibration 355
 - 13.1.2 Image Rectification 356
- Human Vision: Stereopsis 358
- 13.2 Binocular Fusion 362
 - 13.2.1 Correlation 362
 - 13.2.2 Multi-Scale Edge Matching 364
 - 13.2.3 Dynamic Programming 367
- 13.3 Using More Cameras 369
 - 13.3.1 Trinocular Stereo 369
 - 13.3.2 Multiple-Baseline Stereo 371
- 13.4 Notes 372
- 13.5 Assignments 374

14 AFFINE STRUCTURE FROM MOTION 377

- 14.1 Elements of Affine Geometry 378
- 14.2 Affine Structure from Two Images 381
 - 14.2.1 The Affine Structure-from-Motion Theorem 382

- 14.2.2 Rigidity and Metric Constraints 382
- 14.3 Affine Structure from Multiple Images 384
- 14.3.1 The Affine Structure of Affine Image Sequences 385
- Technique: Singular Value Decomposition 385
- 14.3.2 A Factorization Approach to Affine Motion Analysis 387
- 14.4 From Affine to Euclidean Images 388
- 14.4.1 Euclidean Projection Models 389
- 14.4.2 From Affine to Euclidean Motion 390
- 14.5 Affine Motion Segmentation 391
- 14.5.1 The Reduced Echelon Form of the Data Matrix 391
- 14.5.2 The Shape Interaction Matrix 392
- 14.6 Notes 394
- 14.7 Assignments 395

15 PROJECTIVE STRUCTURE FROM MOTION 397

- 15.1 Elements of Projective Geometry 398
- 15.1.1 Projective Bases and Projective Coordinates 398
- 15.1.2 Projective Transformations 400
- 15.1.3 Affine and Projective Spaces 402
- 15.1.4 Hyperplanes and Duality 403
- 15.1.5 Cross-Ratios 404
- 15.1.6 Application: Parameterizing the Fundamental Matrix 407
- 15.2 Projective Scene Reconstruction from Two Views 408
- 15.2.1 Analytical Scene Reconstruction 408
- 15.2.2 Geometric Scene Reconstruction 410
- 15.3 Motion Estimation from Two or Three Views 411
- 15.3.1 Motion Estimation from Fundamental Matrices 412
- 15.3.2 Motion Estimation from Trifocal Tensors 413
- 15.4 Motion Estimation from Multiple Views 415
- 15.4.1 A Factorization Approach to Projective Motion Analysis 415
- 15.4.2 Bundle Adjustment 418
- 15.5 From Projective to Euclidean Structure and Motion 418
- 15.5.1 Metric Upgrades from (Partial) Camera Calibration 419
- 15.5.2 Metric Upgrades from Minimal Assumptions 421
- 15.6 Notes 424
- 15.7 Assignments 426

V MID-LEVEL VISION 431

16 SEGMENTATION BY CLUSTERING 433

- 16.1 What is Segmentation? 433
- 16.1.1 Four Model Problems 435
- 16.1.2 Segmentation as Clustering 436
- 16.2 Human vision: Grouping and Gestalt 437
- 16.3 Applications: Shot Boundary Detection and Background Subtraction 442
- 16.3.1 Background Subtraction 442
- 16.3.2 Shot Boundary Detection 444
- 16.4 Image Segmentation by Clustering Pixels 446
- 16.4.1 Segmentation Using Simple Clustering Methods 447
- 16.4.2 Clustering and Segmentation by K-means 450
- 16.5 Segmentation by Graph-Theoretic Clustering 451
- 16.5.1 Terminology for Graphs 452
- 16.5.2 The Overall Approach 454
- 16.5.3 Affinity Measures 454
- 16.5.4 Eigenvectors and Segmentation 457
- 16.5.5 Normalized Cuts 460
- 16.6 Discussion 463
- 16.6.1 Segmentation and Grouping in People 465
- 16.6.2 Perceptual Grouping 466

17 SEGMENTATION BY FITTING A MODEL 469

- 17.1 Fitting Lines 469
- 17.1.1 The Hough Transform 470

- 17.1.2 Line Fitting with Least Squares 474
- 17.1.3 Which Point is on Which Line? 476
- 17.2 Fitting Curves 478
 - 17.2.1 Implicit Curves 478
 - 17.2.2 Parametric Curves 481
- 17.3 Example: Finding Body Segments by Fitting 482
 - 17.3.1 Some Relations Between Surfaces and Outlines 482
 - 17.3.2 Using Constraints to Fit SOR Outlines 484
- 17.4 Fitting as a Probabilistic Inference Problem 486
- 17.5 Robustness 488
 - 17.5.1 M-estimators 490
 - 17.5.2 RANSAC 492
- 17.6 Example: Using RANSAC to Fit Fundamental Matrices 496
 - 17.6.1 An Expression for Fitting Error 496
 - 17.6.2 Correspondence as Noise 497
 - 17.6.3 Applying RANSAC 497
- 17.7 Discussion 498

18 SEGMENTATION AND FITTING USING PROBABILISTIC METHODS 501

- 18.1 Missing Data Problems, Fitting and Segmentation 502
 - 18.1.1 Missing Data Problems 502
 - 18.1.2 The EM Algorithm 506
 - 18.1.3 The EM Algorithm in the General Case 507
- 18.2 The EM Algorithm in Practice 507
 - 18.2.1 Example: Image Segmentation, Revisited 507
 - 18.2.2 Example: Line Fitting with EM 508
 - 18.2.3 Example: Motion Segmentation and EM 510
 - 18.2.4 Example: Using EM to Identify Outliers 516
 - 18.2.5 Example: Background Subtraction using EM 517
 - 18.2.6 Example: Finding Body Segments with EM 517
 - 18.2.7 Example: EM and the Fundamental Matrix 518
 - 18.2.8 Difficulties with the EM Algorithm 519
- 18.3 How Many are There? 519
 - 18.3.1 Basic Ideas 520
 - 18.3.2 AIC — An Information Criterion 520
 - 18.3.3 Bayesian methods and Schwartz 'BIC 521
 - 18.3.4 Description Length 522
 - 18.3.5 Other Methods for Estimating Deviance 522
- 18.4 Discussion 523
 - 18.4.1 EM and Missing Variable Models 523
 - 18.4.2 Model Selection 524

19 TRACKING WITH LINEAR DYNAMIC MODELS 526

- 19.1 Tracking as an Abstract Inference Problem 527
 - 19.1.1 Independence Assumptions 528
 - 19.1.2 Tracking as Inference 528
 - 19.1.3 Overview 529
- 19.2 Linear Dynamic Models 529
 - 19.2.1 Drifting Points 530
 - 19.2.2 Constant Velocity 530
 - 19.2.3 Constant Acceleration 531
 - 19.2.4 Periodic Motion 531
 - 19.2.5 Higher Order Models 533
- 19.3 Kalman Filtering 534
 - 19.3.1 The Kalman Filter for a 1D State Vector 534
 - 19.3.2 The Kalman Update Equations for a General State Vector 537
 - 19.3.3 Forward-Backward Smoothing 538
- 19.4 Applications and Examples 545
 - 19.4.1 Vehicle Tracking 545
- 19.5 Discussion 549

20 TRACKING WITH NON-LINEAR DYNAMIC MODELS 552

- 20.1 Non-Linear Dynamic Models 552
 - 20.1.1 Unpleasant Properties of Non-Linear Dynamics 553
 - 20.1.2 Difficulties with Likelihoods 554

- 20.2 Particle Filtering 556
 - 20.2.1 Sampled Representations of Probability Distributions 556
 - 20.2.2 The Simplest Particle Filter 560
 - 20.2.3 A Workable Particle Filter 563
 - 20.2.4 If's, And's and But's —Practical Issues in Building Particle Filters 564
- 20.3 Tracking People with Particle Filters 568
- 20.4 Data Association 570
 - 20.4.1 Choosing the Nearest —Global Nearest Neighbours 570
 - 20.4.2 Gating and Probabilistic Data Association 572
- 20.5 Discussion 576
 - 20.5.1 The Particle Filter 576
 - 20.5.2 Starting a People Tracker 577
- I Appendix: The Extended Kalman Filter, or EKF 580

VI APPLICATIONS AND TOPICS 582

21 RANGE DATA 584

- 21.1 Active Range Sensors 584
- 21.2 Range Data Segmentation 587
 - Technique: Analytical Differential Geometry 588
 - 21.2.1 Finding Step and Roof Edges in Range Images 590
 - 21.2.2 Segmenting Range Images into Planar Regions 595
- 21.3 Range Image Registration and Model Construction 597
 - Technique: Quaternions 598
 - 21.3.1 Registering Range Images Using the Iterative Closest-Point Method 599
 - 21.3.2 Fusing Multiple Range Images 602
- 21.4 Object Recognition 603
 - 21.4.1 Matching Piecewise-Planar Surfaces Using Interpretation Trees 604
 - 21.4.2 Matching Free-Form Surfaces Using Spin Images 607
- 21.5 Notes 612
- 21.6 Assignments 613

22 APPLICATION: FINDING IN DIGITAL LIBRARIES 615

- 22.1 Background: Organizing Collections of Information 617
 - 22.1.1 How Well does the System Work? 617
 - 22.1.2 What do Users want? 619
 - 22.1.3 Searching for Pictures 619
 - 22.1.4 Structuring and Browsing 621
- 22.2 Summary Representations of the Whole Picture 622
 - 22.2.1 Histograms and Correlograms 622
 - 22.2.2 Textures and Textures of Textures 624
- 22.3 Representations of Parts of the Picture 627
 - 22.3.1 Segmentation 627
 - 22.3.2 Template matching 628
 - 22.3.3 Shape and correspondence 629
- 22.3.4 Clustering and Organising Collections 630
- 22.4 Video 631
- 22.5 Discussion 632

23 APPLICATION: IMAGE-BASED RENDERING 644

- 23.1 Constructing 3D Models from Image Sequences 645
 - 23.1.1 Scene Modeling from Registered Images 645
 - 23.1.2 Scene Modeling from Unregistered Images 653
- 23.2 Transfer-Based Approaches to Image-Based Rendering 657
 - 23.2.1 Affine View Synthesis 658
 - 23.2.2 Euclidean View Synthesis 661
- 23.3 The Light Field 664
- 23.4 Notes 668
- 23.5 Assignments 670

VII HIGH-LEVEL VISION 672

24 CORRESPONDENCE AND POSE CONSISTENCY 674

- 24.1 Initial Assumptions 674
 - 24.1.1 Obtaining Hypotheses 675
- 24.2 Obtaining Hypotheses by Pose Consistency 676
 - 24.2.1 Pose Consistency for Perspective Cameras 677
 - 24.2.2 Affine and Projective Camera Models 679
 - 24.2.3 Linear Combinations of Models 681
- 24.3 Obtaining Hypotheses by Pose Clustering 682
- 24.4 Obtaining Hypotheses Using Invariants 683
 - 24.4.1 Invariants for Plane Figures 684
 - 24.4.2 Geometric Hashing 689
 - 24.4.3 Invariants and Indexing 689
- 24.5 Verification 694
 - 24.5.1 Edge Proximity 694
 - 24.5.2 Similarity in Texture, Pattern and Intensity 695
 - 24.5.3 Example: Bayes Factors and Verification 696
- 24.6 Application: Registration in Medical Imaging Systems 698
 - 24.6.1 Imaging Modes 698
 - 24.6.2 Applications of Registration 698
 - 24.6.3 Geometric Hashing Techniques in Medical Imaging 700
- 24.7 Curved Surfaces and Alignment 703
- 24.8 Discussion 704
 - 24.8.1 Medical applications 706

25 FINDING TEMPLATES USING CLASSIFIERS 709

- 25.1 Classifiers 710
 - 25.1.1 Using Loss to Determine Decisions 710
 - 25.1.2 Overview: Methods for Building Classifiers 712
 - 25.1.3 Example: A Plug-in Classifier for Normal Class-conditional Densities 714
 - 25.1.4 Example: A Non-Parametric Classifier using Nearest Neighbours 715
 - 25.1.5 Estimating and Improving Performance 716
- 25.2 Building Classifiers from Class Histograms 718
 - 25.2.1 Finding Skin Pixels using a Classifier 719
 - 25.2.2 Face Finding Assuming Independent Template Responses 720
- 25.3 Feature Selection 723
 - 25.3.1 Principal Component Analysis 724
 - 25.3.2 Identifying Individuals with Principal Components Analysis 726
 - 25.3.3 Canonical Variates 729
- 25.4 Neural Networks 733
 - 25.4.1 Key Ideas 733
 - 25.4.2 Minimizing the Error 738
 - 25.4.3 When to Stop Training 740
 - 25.4.4 Finding Faces using Neural Networks 740
 - 25.4.5 Convolutional Neural Nets 741
- 25.5 The Support Vector Machine 743
 - 25.5.1 Support Vector Machines for Linearly Separable Datasets 745
 - 25.5.2 Finding Pedestrians using Support Vector Machines 747
- 25.6 Conclusions 749
 - 25.6.1 Skin Detection 751
 - 25.6.2 Face Finding 752
 - 25.6.3 Pedestrian Finding 752
- I Appendix: Backpropagation 754
- II Appendix: Support Vector Machines for Datasets that are not Linearly Separable 758
- III Appendix: Using Support Vector Machines with Non-Linear Kernels 759

26 RECOGNITION BY RELATIONS BETWEEN TEMPLATES 761

- 26.1 Finding Objects by Voting on Relations between Templates 762
 - 26.1.1 Describing Image Patches 762
 - 26.1.2 Voting and a Simple Generative Model 763
 - 26.1.3 Probabilistic Models for Voting 763
 - 26.1.4 Voting on Relations 766
 - 26.1.5 Voting and 3D Objects 766
- 26.2 Relational Reasoning using Probabilistic Models and Search 767
 - 26.2.1 Correspondence and Search 769

- 26.2.2 Example: Finding Faces 772
- 26.3 Using Classifiers to Prune Search 772
- 26.3.1 Identifying Acceptable Assemblies Using Projected Classifiers 773
- 26.3.2 Example: Finding People and Horses Using Spatial Relations 774
- 26.4 Technique: Hidden Markov Models 776
- 26.4.1 Formal Matters 777
- 26.4.2 Computing with Hidden Markov Models 778
- 26.4.3 Varieties of HMM's 782
- 26.5 Application: Hidden Markov Models and Sign Language Understanding 783
- 26.5.1 Language Models: Sentences from Words 784
- 26.6 Application: Finding People with Hidden Markov Models 786
- 26.7 Conclusions 788
- 26.7.1 Hidden Markov Models 788

27 SMOOTH SURFACES AND THEIR OUTLINES 802

- 27.1 Elements of Differential Geometry 805
- 27.1.1 Curves 805
- 27.1.2 Surfaces 811
- 27.1.3 The Shape of Specularities 815
- 27.2 Contour Geometry 818
- 27.2.1 The Occluding Contour and the Image Contour 819
- 27.2.2 The Cusps and Inflections of the Image Contour 820
- 27.2.3 Koenderink's Theorem 821
- 27.3 Notes 823
- 27.4 Assignments 824

28 ASPECT GRAPHS 826

- 28.1 Differential Geometry and Visual Events 831
- 28.1.1 The Geometry of the Gauss Map 831
- 28.1.2 Asymptotic Curves 833
- 28.1.3 The Asymptotic Spherical Map 835
- 28.1.4 Local Visual Events 836
- 28.1.5 The Bitangent Ray Manifold 838
- 28.1.6 Multilocal Visual Events 840
- 28.1.7 Remarks 841
- 28.2 Computing the Aspect Graph 843
- 28.2.1 Step 1: Tracing Visual Events 844
- 28.2.2 Step 2: Constructing the Regions 845
- 28.2.3 Remaining Steps of the Algorithm 845
- 28.2.4 An Example 847
- 28.3 Aspect Graphs and Object Localization 850
- 28.4 Notes 854
- 28.5 Assignments 856

29 TOWARD CATEGORY-LEVEL OBJECT RECOGNITION 857

- 29.1 The State of the Art and its Limitations 858
- 29.1.1 Current Approaches to Object Recognition 859
- 29.1.2 Limitations 860
- 29.1.3 From Templates to Primitives 861
- 29.1.4 Models of Object Recognition 862
- 29.2 Primitives and Object Recognition 863
- 29.2.1 Volumetric Primitives and Part-Whole Decompositions 863
- 29.2.2 The Two-Dimensional Case: Ribbons 864
- 29.2.3 A Two-Dimensional Recognition System: FORMS 866
- 29.2.4 The Three-Dimensional Case: Generalized Cylinders 870
- 29.2.5 A Three-Dimensional Recognition System: ACRONYM 872
- 29.2.6 Going Further 878
- 29.3 Notes 878
- 29.4 Assignments 879

SUBJECT INDEX 882

BIBLIOGRAPHY 888