

1. Preface
2. Contents
3. Chapter 1 Introduction
4. 1.1 What is computer vision?
5. 1.2 A brief history
6. 1.3 Book overview
7. 1.4 Sample syllabus
8. 1.5 A note on notation
9. 1.6 Additional reading
10. Chapter 2 Image formation
11. 2.1 Geometric primitives and transformations
12. 2.1.1 2D transformations
13. 2.1.2 3D transformations
14. 2.1.3 3D rotations
15. 2.1.4 3D to 2D projections
16. 2.1.5 Lens distortions
17. 2.2 Photometric image formation
18. 2.2.1 Lighting
19. 2.2.2 Reflectance and shading
20. 2.2.3 Optics
21. 2.3 The digital camera
22. 2.3.1 Sampling and aliasing
23. 2.3.2 Color
24. 2.3.3 Compression
25. 2.4 Additional reading
26. 2.5 Exercises
27. Chapter 3 Image processing
28. 3.1 Point operators
29. 3.1.1 Pixel transforms
30. 3.1.2 Color transforms
31. 3.1.3 Compositing and matting
32. 3.1.4 Histogram equalization
33. 3.2 Linear filtering
34. 3.2.1 Separable filtering
35. 3.2.2 Examples of linear filtering
36. 3.2.3 Band-pass and steerable filters
37. 3.3 More neighborhood operators
38. 3.3.1 Non-linear filtering
39. 3.3.2 Bilateral filtering
40. 3.3.3 Binary image processing
41. 3.4 Fourier transforms
42. 3.4.1 Two-dimensional Fourier transforms
43. 3.4.2 Application: Sharpening, blur, and noise removal
44. 3.5 Pyramids and wavelets
45. 3.5.1 Interpolation
46. 3.5.2 Decimation

- 47. 3.5.3 Multi-resolution representations
- 48. 3.5.4 Wavelets
- 49. 3.5.5 Application: Image blending
- 50. 3.6 Geometric transformations
- 51. 3.6.1 Parametric transformations
- 52. 3.6.2 Mesh-based warping
- 53. 3.6.3 Application: Feature-based morphing
- 54. 3.7 Additional reading
- 55. 3.8 Exercises
- 56. Chapter 4 Model fitting and optimization
- 57. 4.1 Scattered data interpolation
- 58. 4.1.1 Radial basis functions
- 59. 4.1.2 Overfitting and underfitting
- 60. 4.1.3 Robust data fitting
- 61. 4.2 Variational methods and regularization
- 62. 4.2.1 Discrete energy minimization
- 63. 4.2.2 Total variation
- 64. 4.2.3 Bilateral solver
- 65. 4.2.4 Application: Interactive colorization
- 66. 4.3 Markov random fields
- 67. 4.3.1 Conditional random fields
- 68. 4.3.2 Application: Interactive segmentation
- 69. 4.4 Additional reading
- 70. 4.5 Exercises
- 71. Chapter 5 Deep Learning
- 72. 5.1 Supervised learning
- 73. 5.1.1 Nearest neighbors
- 74. 5.1.2 Bayesian classification
- 75. 5.1.3 Logistic regression
- 76. 5.1.4 Support vector machines
- 77. 5.1.5 Decision trees and forests
- 78. 5.2 Unsupervised learning
- 79. 5.2.1 Clustering
- 80. 5.2.2 K-means and Gaussians mixture models
- 81. 5.2.3 Principal component analysis
- 82. 5.2.4 Manifold learning
- 83. 5.2.5 Semi-supervised learning
- 84. 5.3 Deep neural networks
- 85. 5.3.1 Weights and layers
- 86. 5.3.2 Activation functions
- 87. 5.3.3 Regularization and normalization
- 88. 5.3.4 Loss functions
- 89. 5.3.5 Backpropagation
- 90. 5.3.6 Training and optimization
- 91. 5.4 Convolutional neural networks
- 92. 5.4.1 Pooling and unpooling

- 93. 5.4.2 Application: Digit classification
- 94. 5.4.3 Network architectures
- 95. 5.4.4 Model zoos
- 96. 5.4.5 Visualizing weights and activations
- 97. 5.4.6 Adversarial examples
- 98. 5.4.7 Self-supervised learning
- 99. 5.5 More complex models
- 100. 5.5.1 Three-dimensional CNNs
- 101. 5.5.2 Recurrent neural networks
- 102. 5.5.3 Transformers
- 103. 5.5.4 Generative models
- 104. 5.6 Additional reading
- 105. 5.7 Exercises
- 106. Chapter 6 Recognition
- 107. 6.1 Instance recognition
- 108. 6.2 Image classification
- 109. 6.2.1 Feature-based methods
- 110. 6.2.2 Deep networks
- 111. 6.2.3 Application: Visual similarity search
- 112. 6.2.4 Face recognition
- 113. 6.3 Object detection
- 114. 6.3.1 Face detection
- 115. 6.3.2 Pedestrian detection
- 116. 6.3.3 General object detection
- 117. 6.4 Semantic segmentation
- 118. 6.4.1 Application: Medical image segmentation
- 119. 6.4.2 Instance segmentation
- 120. 6.4.3 Panoptic segmentation
- 121. 6.4.4 Application: Intelligent photo editing
- 122. 6.4.5 Pose estimation
- 123. 6.5 Video understanding
- 124. 6.6 Vision and language
- 125. 6.7 Additional reading
- 126. 6.8 Exercises
- 127. Chapter 7 Feature detection and matching
- 128. 7.1 Points and patches
- 129. 7.1.1 Feature detectors
- 130. 7.1.2 Feature descriptors
- 131. 7.1.3 Feature matching
- 132. 7.1.4 Large-scale matching and retrieval
- 133. 7.1.5 Feature tracking
- 134. 7.1.6 Application: Performance-driven animation
- 135. 7.2 Edges and contours
- 136. 7.2.1 Edge detection
- 137. 7.2.2 Contour detection
- 138. 7.2.3 Application: Edge editing and enhancement

139.	7.3 Contour tracking
140.	7.3.1 Snakes and scissors
141.	7.3.2 Level Sets
142.	7.3.3 Application: Contour tracking and rotoscoping
143.	7.4 Lines and vanishing points
144.	7.4.1 Successive approximation
145.	7.4.2 Hough transforms
146.	7.4.3 Vanishing points
147.	7.5 Segmentation
148.	7.5.1 Graph-based segmentation
149.	7.5.2 Mean shift
150.	7.5.3 Normalized cuts
151.	7.6 Additional reading
152.	7.7 Exercises
153.	Chapter 8 Image alignment and stitching
154.	8.1 Pairwise alignment
155.	8.1.1 2D alignment using least squares
156.	8.1.2 Application: Panography
157.	8.1.3 Iterative algorithms
158.	8.1.4 Robust least squares and RANSAC
159.	8.1.5 3D alignment
160.	8.2 Image stitching
161.	8.2.1 Parametric motion models
162.	8.2.2 Application: Whiteboard and document scanning
163.	8.2.3 Rotational panoramas
164.	8.2.4 Gap closing
165.	8.2.5 Application: Video summarization and compression
166.	8.2.6 Cylindrical and spherical coordinates
167.	8.3 Global alignment
168.	8.3.1 Bundle adjustment
169.	8.3.2 Parallax removal
170.	8.3.3 Recognizing panoramas
171.	8.4 Compositing
172.	8.4.1 Choosing a compositing surface
173.	8.4.2 Pixel selection and weighting (deghosting)
174.	8.4.3 Application: Photomontage
175.	8.4.4 Blending
176.	8.5 Additional reading
177.	8.6 Exercises
178.	Chapter 9 Motion estimation
179.	9.1 Translational alignment
180.	9.1.1 Hierarchical motion estimation
181.	9.1.2 Fourier-based alignment
182.	9.1.3 Incremental refinement
183.	9.2 Parametric motion
184.	9.2.1 Application: Video stabilization

185.	9.2.2 Spline-based motion
186.	9.2.3 Application: Medical image registration
187.	9.3 Optical flow
188.	9.3.1 Deep learning approaches
189.	9.3.2 Application: Rolling shutter wobble removal
190.	9.3.3 Multi-frame motion estimation
191.	9.3.4 Application: Video denoising
192.	9.4 Layered motion
193.	9.4.1 Application: Frame interpolation
194.	9.4.2 Transparent layers and reflections
195.	9.4.3 Video object segmentation
196.	9.4.4 Video object tracking
197.	9.5 Additional reading
198.	9.6 Exercises
199.	Chapter 10 Computational photography
200.	10.1 Photometric calibration
201.	10.1.1 Radiometric response function
202.	10.1.2 Noise level estimation
203.	10.1.3 Vignetting
204.	10.1.4 Optical blur (spatial response) estimation
205.	10.2 High dynamic range imaging
206.	10.2.1 Tone mapping
207.	10.2.2 Application: Flash photography
208.	10.3 Super-resolution, denoising, and blur removal
209.	10.3.1 Color image demosaicing
210.	10.3.2 Lens blur (bokeh)
211.	10.4 Image matting and compositing
212.	10.4.1 Blue screen matting
213.	10.4.2 Natural image matting
214.	10.4.3 Optimization-based matting
215.	10.4.4 Smoke, shadow, and flash matting
216.	10.4.5 Video matting
217.	10.5 Texture analysis and synthesis
218.	10.5.1 Application: Hole filling and inpainting
219.	10.5.2 Application: Non-photorealistic rendering
220.	10.5.3 Neural style transfer and semantic image synthesis
221.	10.6 Additional reading
222.	10.7 Exercises
223.	Chapter 11 Structure from motion and SLAM
224.	11.1 Geometric intrinsic calibration
225.	11.1.1 Vanishing points
226.	11.1.2 Application: Single view metrology
227.	11.1.3 Rotational motion
228.	11.1.4 Radial distortion
229.	11.2 Pose estimation
230.	11.2.1 Linear algorithms

231.	11.2.2 Iterative non-linear algorithms
232.	11.2.3 Application: Location recognition
233.	11.2.4 Triangulation
234.	11.3 Two-frame structure from motion
235.	11.3.1 Eight, seven, and five-point algorithms
236.	11.3.2 Special motions and structures
237.	11.3.3 Projective (uncalibrated) reconstruction
238.	11.3.4 Self-calibration
239.	11.3.5 Application: View morphing
240.	11.4 Multi-frame structure from motion
241.	11.4.1 Factorization
242.	11.4.2 Bundle adjustment
243.	11.4.3 Exploiting sparsity
244.	11.4.4 Application: Match move
245.	11.4.5 Uncertainty and ambiguities
246.	11.4.6 Application: Reconstruction from internet photos
247.	11.4.7 Global structure from motion
248.	11.4.8 Constrained structure and motion
249.	11.5 Simultaneous localization and mapping (SLAM)
250.	11.5.1 Application: Autonomous navigation
251.	11.5.2 Application: Smartphone augmented reality
252.	11.6 Additional reading
253.	11.7 Exercises
254.	Chapter 12 Depth estimation
255.	12.1 Epipolar geometry
256.	12.1.1 Rectification
257.	12.1.2 Plane sweep
258.	12.2 Sparse correspondence
259.	12.2.1 3D curves and profiles
260.	12.3 Dense correspondence
261.	12.3.1 Similarity measures
262.	12.4 Local methods
263.	12.4.1 Sub-pixel estimation and uncertainty
264.	12.4.2 Application: Stereo-based head tracking
265.	12.5 Global optimization
266.	12.5.1 Dynamic programming
267.	12.5.2 Segmentation-based techniques
268.	12.5.3 Application: Z-keying and background replacement
269.	12.6 Deep neural networks
270.	12.7 Multi-view stereo
271.	12.7.1 Scene flow
272.	12.7.2 Volumetric and 3D surface reconstruction
273.	12.7.3 Shape from silhouettes
274.	12.8 Monocular depth estimation
275.	12.9 Additional reading
276.	12.10 Exercises

277.	Chapter 13 3D reconstruction
278.	13.1 Shape from X
279.	13.1.1 Shape from shading and photometric stereo
280.	13.1.2 Shape from texture
281.	13.1.3 Shape from focus
282.	13.2 3D scanning
283.	13.2.1 Range data merging
284.	13.2.2 Application: Digital heritage
285.	13.3 Surface representations
286.	13.3.1 Surface interpolation
287.	13.3.2 Surface simplification
288.	13.3.3 Geometry images
289.	13.4 Point-based representations
290.	13.5 Volumetric representations
291.	13.5.1 Implicit surfaces and level sets
292.	13.6 Model-based reconstruction
293.	13.6.1 Architecture
294.	13.6.2 Facial modeling and tracking
295.	13.6.3 Application: Facial animation
296.	13.6.4 Human body modeling and tracking
297.	13.7 Recovering texture maps and albedos
298.	13.7.1 Estimating BRDFs
299.	13.7.2 Application: 3D model capture
300.	13.8 Additional reading
301.	13.9 Exercises
302.	Chapter 14 Image-based rendering
303.	14.1 View interpolation
304.	14.1.1 View-dependent texture maps
305.	14.1.2 Application: Photo Tourism
306.	14.2 Layered depth images
307.	14.2.1 Impostors, sprites, and layers
308.	14.2.2 Application: 3D photography
309.	14.3 Light fields and Lumigraphs
310.	14.3.1 Unstructured Lumigraph
311.	14.3.2 Surface light fields
312.	14.3.3 Application: Concentric mosaics
313.	14.3.4 Application: Synthetic re-focusing
314.	14.4 Environment mattes
315.	14.4.1 Higher-dimensional light fields
316.	14.4.2 The modeling to rendering continuum
317.	14.5 Video-based rendering
318.	14.5.1 Video-based animation
319.	14.5.2 Video textures
320.	14.5.3 Application: Animating pictures
321.	14.5.4 3D and free-viewpoint Video
322.	14.5.5 Application: Video-based walkthroughs

323.	14.6 Neural rendering
324.	14.7 Additional reading
325.	14.8 Exercises
326.	Chapter 15 Conclusion
327.	Appendix A Linear algebra and numerical techniques
328.	A.1 Matrix decompositions
329.	A.1.1 Singular value decomposition
330.	A.1.2 Eigenvalue decomposition
331.	A.1.3 QR factorization
332.	A.1.4 Cholesky factorization
333.	A.2 Linear least squares
334.	A.2.1 Total least squares
335.	A.3 Non-linear least squares
336.	A.4 Direct sparse matrix techniques
337.	A.4.1 Variable reordering
338.	A.5 Iterative techniques
339.	A.5.1 Conjugate gradient
340.	A.5.2 Preconditioning
341.	A.5.3 Multigrid
342.	Appendix B Bayesian modeling and inference
343.	B.1 Estimation theory
344.	B.2 Maximum likelihood estimation and least squares
345.	B.3 Robust statistics
346.	B.4 Prior models and Bayesian inference
347.	B.5 Markov random fields
348.	B.6 Uncertainty estimation (error analysis)
349.	Appendix C Supplementary material
350.	C.1 Datasets and benchmarks
351.	C.2 Software
352.	C.3 Slides and lectures
353.	References
354.	Index