Contents

Preface ix

1 Introduction 1
- 1.1 What is computer vision? . . . 3
- 1.2 A brief history . . . 10
- 1.3 Book overview . . . 17
- 1.4 Sample syllabus . . . 23
- 1.5 A note on notation . . . 25
- 1.6 Additional reading . . . 25

2 Image formation 27
- 2.1 Geometric primitives and transformations . . . 29
  - 2.1.1 Geometric primitives . . . 29
  - 2.1.2 2D transformations . . . 33
  - 2.1.3 3D transformations . . . 36
  - 2.1.4 3D rotations . . . 37
  - 2.1.5 3D to 2D projections . . . 42
  - 2.1.6 Lens distortions . . . 52
- 2.2 Photometric image formation . . . 54
  - 2.2.1 Lighting . . . 54
  - 2.2.2 Reflectance and shading . . . 55
  - 2.2.3 Optics . . . 61
- 2.3 The digital camera . . . 65
  - 2.3.1 Sampling and aliasing . . . 69
  - 2.3.2 Color . . . 71
  - 2.3.3 Compression . . . 80
- 2.4 Additional reading . . . 82
- 2.5 Exercises . . . 82

3 Image processing 87
- 3.1 Point operators . . . 89
  - 3.1.1 Pixel transforms . . . 91
  - 3.1.2 Color transforms . . . 92
  - 3.1.3 Compositing and matting . . . 92
  - 3.1.4 Histogram equalization . . . 94

3.1.5 *Application*: Tonal adjustment 97
3.2 Linear filtering 98
3.2.1 Separable filtering 102
3.2.2 Examples of linear filtering 103
3.2.3 Band-pass and steerable filters 104
3.3 More neighborhood operators 108
3.3.1 Non-linear filtering 108
3.3.2 Morphology 112
3.3.3 Distance transforms 113
3.3.4 Connected components 115
3.4 Fourier transforms 116
3.4.1 Fourier transform pairs 119
3.4.2 Two-dimensional Fourier transforms 123
3.4.3 Wiener filtering 123
3.4.4 *Application*: Sharpening, blur, and noise removal 126
3.5 Pyramids and wavelets 127
3.5.1 Interpolation 127
3.5.2 Decimation 130
3.5.3 Multi-resolution representations 132
3.5.4 Wavelets 136
3.5.5 *Application*: Image blending 140
3.6 Geometric transformations 143
3.6.1 Parametric transformations 145
3.6.2 Mesh-based warping 149
3.6.3 *Application*: Feature-based morphing 152
3.7 Global optimization 153
3.7.1 Regularization 154
3.7.2 Markov random fields 158
3.7.3 *Application*: Image restoration 169
3.8 Additional reading 169
3.9 Exercises 171

4 Feature detection and matching **181**
4.1 Points and patches 183
4.1.1 Feature detectors 185
4.1.2 Feature descriptors 196
4.1.3 Feature matching 200
4.1.4 Feature tracking 207
4.1.5 *Application*: Performance-driven animation 209
4.2 Edges 210
4.2.1 Edge detection 210
4.2.2 Edge linking 215
4.2.3 *Application*: Edge editing and enhancement 219
4.3 Lines 220
4.3.1 Successive approximation 220
4.3.2 Hough transforms 221

4.3.3 Vanishing points . . . 224
4.3.4 *Application*: Rectangle detection . . . 226
4.4 Additional reading . . . 227
4.5 Exercises . . . 228

5 Segmentation **235**
5.1 Active contours . . . 237
5.1.1 Snakes . . . 238
5.1.2 Dynamic snakes and CONDENSATION . . . 243
5.1.3 Scissors . . . 246
5.1.4 Level Sets . . . 248
5.1.5 *Application*: Contour tracking and rotoscoping . . . 249
5.2 Split and merge . . . 250
5.2.1 Watershed . . . 251
5.2.2 Region splitting (divisive clustering) . . . 251
5.2.3 Region merging (agglomerative clustering) . . . 251
5.2.4 Graph-based segmentation . . . 252
5.2.5 Probabilistic aggregation . . . 253
5.3 Mean shift and mode finding . . . 254
5.3.1 K-means and mixtures of Gaussians . . . 256
5.3.2 Mean shift . . . 257
5.4 Normalized cuts . . . 260
5.5 Graph cuts and energy-based methods . . . 264
5.5.1 *Application*: Medical image segmentation . . . 268
5.6 Additional reading . . . 268
5.7 Exercises . . . 270

6 Feature-based alignment **273**
6.1 2D and 3D feature-based alignment . . . 275
6.1.1 2D alignment using least squares . . . 275
6.1.2 *Application*: Panography . . . 277
6.1.3 Iterative algorithms . . . 278
6.1.4 Robust least squares and RANSAC . . . 281
6.1.5 3D alignment . . . 283
6.2 Pose estimation . . . 284
6.2.1 Linear algorithms . . . 284
6.2.2 Iterative algorithms . . . 286
6.2.3 *Application*: Augmented reality . . . 287
6.3 Geometric intrinsic calibration . . . 288
6.3.1 Calibration patterns . . . 289
6.3.2 Vanishing points . . . 290
6.3.3 *Application*: Single view metrology . . . 292
6.3.4 Rotational motion . . . 293
6.3.5 Radial distortion . . . 295
6.4 Additional reading . . . 296
6.5 Exercises . . . 296

7 Structure from motion **303**
7.1 Triangulation . . . 305
7.2 Two-frame structure from motion . . . 307
7.2.1 Projective (uncalibrated) reconstruction . . . 312
7.2.2 Self-calibration . . . 313
7.2.3 *Application*: View morphing . . . 315
7.3 Factorization . . . 315
7.3.1 Perspective and projective factorization . . . 318
7.3.2 *Application*: Sparse 3D model extraction . . . 319
7.4 Bundle adjustment . . . 320
7.4.1 Exploiting sparsity . . . 322
7.4.2 *Application*: Match move and augmented reality . . . 324
7.4.3 Uncertainty and ambiguities . . . 326
7.4.4 *Application*: Reconstruction from Internet photos . . . 327
7.5 Constrained structure and motion . . . 329
7.5.1 Line-based techniques . . . 330
7.5.2 Plane-based techniques . . . 331
7.6 Additional reading . . . 332
7.7 Exercises . . . 332

8 Dense motion estimation **335**
8.1 Translational alignment . . . 337
8.1.1 Hierarchical motion estimation . . . 341
8.1.2 Fourier-based alignment . . . 341
8.1.3 Incremental refinement . . . 345
8.2 Parametric motion . . . 350
8.2.1 *Application*: Video stabilization . . . 354
8.2.2 Learned motion models . . . 354
8.3 Spline-based motion . . . 355
8.3.1 *Application*: Medical image registration . . . 358
8.4 Optical flow . . . 360
8.4.1 Multi-frame motion estimation . . . 363
8.4.2 *Application*: Video denoising . . . 364
8.4.3 *Application*: De-interlacing . . . 364
8.5 Layered motion . . . 365
8.5.1 *Application*: Frame interpolation . . . 368
8.5.2 Transparent layers and reflections . . . 368
8.6 Additional reading . . . 370
8.7 Exercises . . . 371

9 Image stitching **375**
9.1 Motion models . . . 378
9.1.1 Planar perspective motion . . . 379
9.1.2 *Application*: Whiteboard and document scanning . . . 379
9.1.3 Rotational panoramas . . . 380
9.1.4 Gap closing . . . 382

9.1.5 *Application*: Video summarization and compression 383
9.1.6 Cylindrical and spherical coordinates 385
9.2 Global alignment 387
9.2.1 Bundle adjustment 388
9.2.2 Parallax removal 391
9.2.3 Recognizing panoramas 392
9.2.4 Direct vs. feature-based alignment 393
9.3 Compositing 396
9.3.1 Choosing a compositing surface 396
9.3.2 Pixel selection and weighting (de-ghosting) 398
9.3.3 *Application*: Photomontage 403
9.3.4 Blending 403
9.4 Additional reading 406
9.5 Exercises 407

10 Computational photography **409**
10.1 Photometric calibration 412
10.1.1 Radiometric response function 412
10.1.2 Noise level estimation 415
10.1.3 Vignetting 416
10.1.4 Optical blur (spatial response) estimation 416
10.2 High dynamic range imaging 419
10.2.1 Tone mapping 427
10.2.2 *Application*: Flash photography 434
10.3 Super-resolution and blur removal 436
10.3.1 Color image demosaicing 440
10.3.2 *Application*: Colorization 442
10.4 Image matting and compositing 443
10.4.1 Blue screen matting 445
10.4.2 Natural image matting 446
10.4.3 Optimization-based matting 450
10.4.4 Smoke, shadow, and flash matting 452
10.4.5 Video matting 454
10.5 Texture analysis and synthesis 455
10.5.1 *Application*: Hole filling and inpainting 457
10.5.2 *Application*: Non-photorealistic rendering 458
10.6 Additional reading 460
10.7 Exercises 461

11 Stereo correspondence **467**
11.1 Epipolar geometry 471
11.1.1 Rectification 472
11.1.2 Plane sweep 474
11.2 Sparse correspondence 475
11.2.1 3D curves and profiles 476
11.3 Dense correspondence 477

11.3.1 Similarity measures . . . 479
11.4 Local methods . . . 480
11.4.1 Sub-pixel estimation and uncertainty . . . 482
11.4.2 *Application*: Stereo-based head tracking . . . 483
11.5 Global optimization . . . 484
11.5.1 Dynamic programming . . . 485
11.5.2 Segmentation-based techniques . . . 487
11.5.3 *Application*: Z-keying and background replacement . . . 489
11.6 Multi-view stereo . . . 489
11.6.1 Volumetric and 3D surface reconstruction . . . 492
11.6.2 Shape from silhouettes . . . 497
11.7 Additional reading . . . 499
11.8 Exercises . . . 500

12 3D reconstruction **505**
12.1 Shape from X . . . 508
12.1.1 Shape from shading and photometric stereo . . . 508
12.1.2 Shape from texture . . . 510
12.1.3 Shape from focus . . . 511
12.2 Active rangefinding . . . 512
12.2.1 Range data merging . . . 515
12.2.2 *Application*: Digital heritage . . . 517
12.3 Surface representations . . . 518
12.3.1 Surface interpolation . . . 518
12.3.2 Surface simplification . . . 520
12.3.3 Geometry images . . . 520
12.4 Point-based representations . . . 521
12.5 Volumetric representations . . . 522
12.5.1 Implicit surfaces and level sets . . . 522
12.6 Model-based reconstruction . . . 523
12.6.1 Architecture . . . 524
12.6.2 Heads and faces . . . 526
12.6.3 *Application*: Facial animation . . . 528
12.6.4 Whole body modeling and tracking . . . 530
12.7 Recovering texture maps and albedos . . . 534
12.7.1 Estimating BRDFs . . . 536
12.7.2 *Application*: 3D photography . . . 537
12.8 Additional reading . . . 538
12.9 Exercises . . . 539

13 Image-based rendering **543**
13.1 View interpolation . . . 545
13.1.1 View-dependent texture maps . . . 547
13.1.2 *Application*: Photo Tourism . . . 548
13.2 Layered depth images . . . 549
13.2.1 Impostors, sprites, and layers . . . 549

13.3 Light fields and Lumigraphs 551
13.3.1 Unstructured Lumigraph 554
13.3.2 Surface light fields 555
13.3.3 *Application*: Concentric mosaics 556
13.4 Environment mattes 556
13.4.1 Higher-dimensional light fields 558
13.4.2 The modeling to rendering continuum 559
13.5 Video-based rendering 560
13.5.1 Video-based animation 560
13.5.2 Video textures 561
13.5.3 *Application*: Animating pictures 564
13.5.4 3D Video 564
13.5.5 *Application*: Video-based walkthroughs 566
13.6 Additional reading 569
13.7 Exercises 570

14 Recognition **575**
14.1 Object detection 578
14.1.1 Face detection 578
14.1.2 Pedestrian detection 585
14.2 Face recognition 588
14.2.1 Eigenfaces 589
14.2.2 Active appearance and 3D shape models 596
14.2.3 *Application*: Personal photo collections 601
14.3 Instance recognition 602
14.3.1 Geometric alignment 603
14.3.2 Large databases 604
14.3.3 *Application*: Location recognition 609
14.4 Category recognition 611
14.4.1 Bag of words 612
14.4.2 Part-based models 615
14.4.3 Recognition with segmentation 620
14.4.4 *Application*: Intelligent photo editing 621
14.5 Context and scene understanding 625
14.5.1 Learning and large image collections 627
14.5.2 *Application*: Image search 630
14.6 Recognition databases and test sets 631
14.7 Additional reading 631
14.8 Exercises 637

15 Conclusion **641**

A Linear algebra and numerical techniques **645**
A.1 Matrix decompositions 646
A.1.1 Singular value decomposition 646
A.1.2 Eigenvalue decomposition 647

A.1.3 QR factorization . . . 649
A.1.4 Cholesky factorization . . . 650
A.2 Linear least squares . . . 651
A.2.1 Total least squares . . . 653
A.3 Non-linear least squares . . . 654
A.4 Direct sparse matrix techniques . . . 655
A.4.1 Variable reordering . . . 656
A.5 Iterative techniques . . . 656
A.5.1 Conjugate gradient . . . 657
A.5.2 Preconditioning . . . 659
A.5.3 Multigrid . . . 660

B Bayesian modeling and inference **661**
B.1 Estimation theory . . . 662
B.1.1 Likelihood for multivariate Gaussian noise . . . 663
B.2 Maximum likelihood estimation and least squares . . . 665
B.3 Robust statistics . . . 666
B.4 Prior models and Bayesian inference . . . 667
B.5 Markov random fields . . . 668
B.5.1 Gradient descent and simulated annealing . . . 670
B.5.2 Dynamic programming . . . 670
B.5.3 Belief propagation . . . 672
B.5.4 Graph cuts . . . 674
B.5.5 Linear programming . . . 676
B.6 Uncertainty estimation (error analysis) . . . 678

C Supplementary material **679**
C.1 Data sets . . . 680
C.2 Software . . . 682
C.3 Slides and lectures . . . 689
C.4 Bibliography . . . 690

References **691**

Index **793**