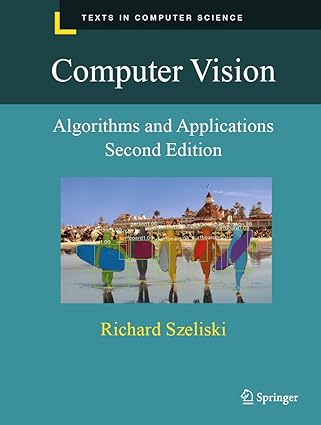
Computer Vision - Algorithms and Applications

*Springer – 2022*



# 1 Introduction

1.1 What is computer vision

1.2 A brief history

1.3 Book overview

1.4 Sample syllabus

1.5 A note on notation

1.6 Additional reading

# 2 Image formation

## 2.1 Geometric primitives and transformations

### 2.1.1 2D transformations

### 2.1.2 3D transformations

### 2.1.3 3D rotations

### 2.1.4 3D to 2D projections

### 2.1.5 Lens distortions

## 2.2 Photometric image formation

2.2.1 Lighting

2.2.2 Reflectance and shading

2.2.3 Optics

## 2.3 The digital camera

2.3.1 Sampling and aliasing

2.3.2 Color

2.3.3 Compression

## 2.4 Additional reading

## 2.5 Exercises

# 3 Image processing

3.1 Point operators

3.1.1 Pixel transforms

3.1.2 Color transforms

3.1.3 Compositing and matting

3.1.4 Histogram equalization

3.1.5 Application: Tonal adjustment

3.2 Linear filtering

3.2.1 Separable filtering

3.2.2 Examples of linear filtering

3.2.3 Band-pass and steerable filters

3.3 More neighborhood operators

3.3.1 Non-linear filtering

3.3.2 Bilateral filtering

3.3.3 Binary image processing

3.4 Fourier transforms

3.4.1 Two-dimensional Fourier transforms

3.4.2 Application: Sharpening, blur, and noise removal

3.5 Pyramids and wavelets

3.5.1 Interpolation

3.5.2 Decimation

3.5.3 Multi-resolution representations

3.5.4 Wavelets

3.5.5 Application: Image blending

3.6 Geometric transformations

3.6.1 Parametric transformations

3.6.2 Mesh-based warping

3.6.3 Application: Feature-based morphing

3.7 Additional reading

3.8 Exercises

# 4 Model fitting and optimization

4.1 Scattered data interpolation

4.1.1 Radial basis functions

4.1.2 Overfitting and underfitting

4.1.3 Robust data fitting

4.2 Variational methods and regularization

4.2.1 Discrete energy minimization

4.2.2 Total variation

4.2.3 Bilateral solver

4.2.4 Application: Interactive colorization

4.3 Markov random fields

4.3.1 Conditional random fields

4.3.2 Application: Interactive segmentation

4.4 Additional reading

4.5 Exercises

# 5 Deep Learning

5.1 Supervised learning

5.1.1 Nearest neighbors

5.1.2 Bayesian classification

5.1.3 Logistic regression

5.1.4 Support vector machines

5.1.5 Decision trees and forests

5.2 Unsupervised learning

5.2.1 Clustering

5.2.2 K-means and Gaussians mixture models

5.2.3 Principal component analysis

5.2.4 Manifold learning

5.2.5 Semi-supervised learning

5.3 Deep neural networks

5.3.1 Weights and layers

5.3.2 Activation functions

5.3.3 Regularization and normalization

5.3.4 Loss functions

5.3.5 Backpropagation

5.3.6 Training and optimization

5.4 Convolutional neural networks

5.4.1 Pooling and unpooling

5.4.2 Application: Digit classification

5.4.3 Network architectures

5.4.4 Model zoos

5.4.5 Visualizing weights and activations

5.4.6 Adversarial examples

5.4.7 Self-supervised learning

5.5 More complex models

5.5.1 Three-dimensional CNNs

5.5.2 Recurrent neural networks

5.5.3 Transformers

5.5.4 Generative models

5.6 Additional reading

5.7 Exercises

# 6 Recognition

6.1 Instance recognition

6.2 Image classification

6.2.1 Feature-based methods

6.2.2 Deep networks

6.2.3 Application: Visual similarity search

6.2.4 Face recognition

6.3 Object detection

6.3.1 Face detection

6.3.2 Pedestrian detection

6.3.3 General object detection

6.4 Semantic segmentation

6.4.1 Application: Medical image segmentation

6.4.2 Instance segmentation

6.4.3 Panoptic segmentation

6.4.4 Application: Intelligent photo editing

6.4.5 Pose estimation

6.5 Video understanding

6.6 Vision and language

6.7 Additional reading

6.8 Exercises

# 7 Feature detection and matching

7.1 Points and patches

7.1.1 Feature detectors

7.1.2 Feature descriptors

7.1.3 Feature matching

7.1.4 Large-scale matching and retrieval

7.1.5 Feature tracking

7.1.6 Application: Performance-driven animation

7.2 Edges and contours

7.2.1 Edge detection

7.2.2 Contour detection

7.2.3 Application: Edge editing and enhancement

7.3 Contour tracking

7.3.1 Snakes and scissors

7.3.2 Level Sets

7.3.3 Application: Contour tracking and rotoscoping

7.4 Lines and vanishing points

7.4.1 Successive approximation

7.4.2 Hough transforms

7.4.3 Vanishing points

7.5 Segmentation

7.5.1 Graph-based segmentation

7.5.2 Mean shift

7.5.3 Normalized cuts

7.6 Additional reading

7.7 Exercises

# 8 Image alignment and stitching

8.1 Pairwise alignment

8.1.1 2D alignment using least squares

8.1.2 Application: Panography

8.1.3 Iterative algorithms

8.1.4 Robust least squares and RANSAC

8.1.5 3D alignment

8.2 Image stitching

8.2.1 Parametric motion models

8.2.2 Application: Whiteboard and document scanning

8.2.3 Rotational panoramas

8.2.4 Gap closing

8.2.5 Application: Video summarization and compression

8.2.6 Cylindrical and spherical coordinates

8.3 Global alignment

8.3.1 Bundle adjustment

8.3.2 Parallax removal

8.3.3 Recognizing panoramas

8.4 Compositing

8.4.1 Choosing a compositing surface

8.4.2 Pixel selection and weighting (deghosting)

8.4.3 Application: Photomontage

8.4.4 Blending

8.5 Additional reading

8.6 Exercises

# 9 Motion estimation

9.1 Translational alignment

9.1.1 Hierarchical motion estimation

9.1.2 Fourier-based alignment

9.1.3 Incremental refinement

9.2 Parametric motion

9.2.1 Application: Video stabilization

9.2.2 Spline-based motion

9.2.3 Application: Medical image registration

9.3 Optical flow

9.3.1 Deep learning approaches

9.3.2 Application: Rolling shutter wobble removal

9.3.3 Multi-frame motion estimation

9.3.4 Application: Video denoising

9.4 Layered motion

9.4.1 Application: Frame interpolation

9.4.2 Transparent layers and reflections

9.4.3 Video object segmentation

9.4.4 Video object tracking

9.5 Additional reading

9.6 Exercises

# 10 Computational photography

10.1 Photometric calibration

10.1.1 Radiometric response function

10.1.2 Noise level estimation

10.1.3 Vignetting

10.1.4 Optical blur (spatial response) estimation

10.2 High dynamic range imaging

10.2.1 Tone mapping

10.2.2 Application: Flash photography

10.3 Super-resolution, denoising, and blur removal

10.3.1 Color image demosaicing

10.3.2 Lens blur (bokeh)

10.4 Image matting and compositing

10.4.1 Blue screen matting

10.4.2 Natural image matting

10.4.3 Optimization-based matting

10.4.4 Smoke, shadow, and flash matting

10.4.5 Video matting

10.5 Texture analysis and synthesis

10.5.1 Application: Hole filling and inpainting

10.5.2 Application: Non-photorealistic rendering

10.5.3 Neural style transfer and semantic image synthesis

10.6 Additional reading

10.7 Exercises

# 11 Structure from motion and SLAM

11.1 Geometric intrinsic calibration

11.1.1 Vanishing points

11.1.2 Application: Single view metrology

11.1.3 Rotational motion

11.1.4 Radial distortion

11.2 Pose estimation

11.2.1 Linear algorithms

11.2.2 Iterative non-linear algorithms

11.2.3 Application: Location recognition

11.2.4 Triangulation

11.3 Two-frame structure from motion

11.3.1 Eight, seven, and five-point algorithms

11.3.2 Special motions and structures

11.3.3 Projective (uncalibrated) reconstruction

11.3.4 Self-calibration

11.3.5 Application: View morphing

11.4 Multi-frame structure from motion

11.4.1 Factorization

11.4.2 Bundle adjustment

11.4.3 Exploiting sparsity

11.4.4 Application: Match move

11.4.5 Uncertainty and ambiguities

11.4.6 Application: Reconstruction from internet photos

11.4.7 Global structure from motion

11.4.8 Constrained structure and motion

11.5 Simultaneous localization and mapping (SLAM)

11.5.1 Application: Autonomous navigation

11.5.2 Application: Smartphone augmented reality

11.6 Additional reading

11.7 Exercises

# 12 Depth estimation

12.1 Epipolar geometry

12.1.1 Rectification

12.1.2 Plane sweep

12.2 Sparse correspondence

12.2.1 3D curves and profiles

12.3 Dense correspondence

12.3.1 Similarity measures

12.4 Local methods

12.4.1 Sub-pixel estimation and uncertainty

12.4.2 Application: Stereo-based head tracking

12.5 Global optimization

12.5.1 Dynamic programming

12.5.2 Segmentation-based techniques

12.5.3 Application: Z-keying and background replacement

12.6 Deep neural networks

12.7 Multi-view stereo

12.7.1 Scene flow

12.7.2 Volumetric and 3D surface reconstruction

12.7.3 Shape from silhouettes

12.8 Monocular depth estimation

12.9 Additional reading

12.10Exercises

# 13 3D reconstruction

13.1 Shape from X

13.1.1 Shape from shading and photometric stereo

13.1.2 Shape from texture

13.1.3 Shape from focus

13.2 3D scanning

13.2.1 Range data merging

13.2.2 Application: Digital heritage

13.3 Surface representations

13.3.1 Surface interpolation

13.3.2 Surface simplification

13.3.3 Geometry images

13.4 Point-based representations

13.5 Volumetric representations

13.5.1 Implicit surfaces and level sets

13.6 Model-based reconstruction

13.6.1 Architecture

13.6.2 Facial modeling and tracking

13.6.3 Application: Facial animation

13.6.4 Human body modeling and tracking

13.7 Recovering texture maps and albedos

13.7.1 Estimating BRDFs

13.7.2 Application: 3D model capture

13.8 Additional reading

13.9 Exercises

# 14 Image-based rendering

14.1 View interpolation

14.1.1 View-dependent texture maps

14.1.2 Application: Photo Tourism

14.2 Layered depth images

14.2.1 Impostors, sprites, and layers

14.2.2 Application: 3D photography

14.3 Light fields and Lumigraphs

14.3.1 Unstructured Lumigraph

14.3.2 Surface light fields

14.3.3 Application: Concentric mosaics

14.3.4 Application: Synthetic re-focusing

14.4 Environment mattes

14.4.1 Higher-dimensional light fields

14.4.2 The modeling to rendering continuum

14.5 Video-based rendering

14.5.1 Video-based animation

14.5.2 Video textures

14.5.3 Application: Animating pictures

14.5.4 3D and free-viewpoint Video

14.5.5 Application: Video-based walkthroughs

14.6 Neural rendering

14.7 Additional reading

14.8 Exercises