

Feature Extraction and Image Processing for Computer Vision

5th Edition

Contents

Preface

- What is New in the Fifth Edition?
- Why Did We Write this Book?
- The Book and its Support
- In Gratitude
- Final Message
- Dedication

1 Introduction

- 1.1 Overview
- 1.2 Human and Computer Vision
- 1.3 The Human Vision System
 - 1.3.1 The Eye
 - 1.3.2 The Neural System
 - 1.3.3 Processing
- 1.4 Computer Vision Systems
 - 1.4.1 Cameras
 - 1.4.2 Computer Interfaces
- 1.5 Processing Images
 - 1.5.1 The Deep Learning Revolution
 - 1.5.2 Processing
 - 1.5.3 Hello Python, Hello Images!
 - 1.5.4 Mathematical Tools
 - 1.5.5 Hello Matlab
- 1.6 Associated Literature
 - 1.6.1 Journals, Magazines and Conferences
 - 1.6.2 Textbooks
 - 1.6.3 The Web
- 1.7 Conclusions
- 1.8 Chapter 1 References

2 Images, Sampling and Frequency Domain Processing

- 2.1 Overview
- 2.2 Image Formation
- 2.3 The Fourier Transform
- 2.4 The Sampling Criterion

2.5 The Discrete Fourier Transform (DFT)

- 2.5.1 One Dimensional Transform
- 2.5.2 Two Dimensional Transform

2.6 Properties of the Fourier Transform

- 2.6.1 Shift Invariance
- 2.6.2 Rotation
- 2.6.3 Frequency Scaling
- 2.6.4 Superposition (Linearity)
- 2.6.5 The Importance of Phase

2.7 Transforms other than Fourier

- 2.7.1 Discrete Cosine Transform
- 2.7.2 Discrete Hartley Transform
- 2.7.3 Introductory Wavelets
 - 2.7.3.1 Gabor Wavelet
 - 2.7.3.2 Haar Wavelet
- 2.7.4 Deep Learning based Compression
- 2.7.5 Other Transforms

2.8 Applications using Frequency Domain Properties

- 2.9 Further Reading
- 2.10 Chapter 2 References

3 Image Processing

- 3.1 Overview
- 3.2 Histograms
- 3.3 Point Operators
 - 3.3.1 Basic Point Operations
 - 3.3.2 Histogram Normalisation
 - 3.3.3 Histogram Equalisation
 - 3.3.4 Thresholding
- 3.4 Group Operations
 - 3.4.1 Template Convolution
 - 3.4.2 Averaging Operator
 - 3.4.3 On Different Template Size
 - 3.4.4 Template Convolution via the Fourier Transform
 - 3.4.5 Gaussian Averaging Operator
 - 3.4.6 More on Averaging

- 3.5 Other Image Processing Operators
 - 3.5.1 Median Filter
 - 3.5.2 Mode Filter
- 3.6 Mathematical Morphology
 - 3.6.1 Morphological Operators
 - 3.6.2 Grey Level Morphology
 - 3.6.3 Grey Level Erosion and Dilation
 - 3.6.4 Minkowski Operators
- 3.7 Further Reading
- 3.8 Chapter 3 References

4 Distance, Classification and Deep Learning

- 4.1 Overview
- 4.2 Basis of Classification and Learning
- 4.3 Distance Measures
 - 4.3.1 Basic properties
 - 4.3.2 Basic measures: Manhattan and Euclidean, L_n norms
 - 4.3.3 Measuring groups: Mahalanobis, Bhattacharyya and Matusita
 - 4.3.4 Other measures: Histogram Intersection, Chi2 (χ^2) and the Earth Mover's Distance
- 4.4 Classification
 - 4.4.1 The k-Nearest Neighbour for Classification
 - 4.4.2 Support Vector Machines
 - 4.4.3 Random Forests
 - 4.4.4 Other Approaches
- 4.5 Neural Networks
 - 4.5.1 Basis of Learning: Error Back Propagation
 - 4.5.2 Neuron Operation
 - 4.5.3 Activation Functions
 - 4.5.4 Analysing a Basic MLP
 - 4.5.5 More Activation Functions
 - 4.5.6 Basic Loss Functions
- 4.6 Deep Learning
 - 4.6.1 Basis of Deep Learning
 - 4.6.2 The Beginning of the Revolution
 - 4.6.2.1 Convolutional Layers
 - 4.6.2.2 Pooling Layers
 - 4.6.2.3 Fully Connected Layers
 - 4.6.2.4 Full Architecture
 - 4.6.2.5 Performance

- 4.6.3 Deep Learning Architectures/Backbones and Objectives
 - 4.6.3.1 Resnet, VGG and other Convolutional Networks
 - 4.6.3.2 U-net
 - 4.6.3.3 Applying Deep Learning
- 4.6.4 Building Blocks: Architectures and Strategies
 - 4.6.4.1 Range of Approaches and Performance
 - 4.6.4.2 Feature Pyramid Networks
 - 4.6.4.3 Non-Local Operations
 - 4.6.4.4 Encoder and Autoencoder
 - 4.6.4.5 Generative Adversarial Networks (GAN)
 - 4.6.4.6 Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM)
 - 4.6.4.7 Transformer
 - 4.6.4.8 Diffusion Models
 - 4.6.4.9 Other Components: CapsNet and 3D CNN
 - 4.6.4.10 Recent Developments: Neural Radiance Fields (NeRFs) and Foundation Models
 - 4.6.4.11 Example Architecture
- 4.6.5 Applying Deep Learning
 - 4.6.5.1 Training and Application
 - 4.6.5.2 More Loss Functions
 - 4.6.5.3 Advanced Loss Functions – Arcface and Cosface
- 4.6.6 Data
- 4.6.7 Understanding the Results
- 4.6.8 Applying Deep Learning: Classification
- 4.7 Image Segmentation with Deep Learning
 - 4.7.1 Segmentation by Encoding and Decoding
 - 4.7.2 Region with CNN Features
 - 4.7.3 You Only Look Once
- 4.8 Further Reading
- 4.9 Chapter 4 References

5 Low-Level Feature Extraction (including Edge Detection)

5.1 Overview

5.2 Edge Detection

5.2.1 Handcrafted First Order Edge Detection Operators

5.2.1.1 Basic Operators

5.2.1.2 Analysis of the Basic Operators

5.2.1.3 Prewitt Edge Detection Operator

5.2.1.4 Sobel Edge Detection Operator

5.2.1.5 The Canny Edge Detector

5.2.2 Second Order Edge Detection Operators

5.2.2.1 Motivation

5.2.2.2 Basic Operators: The Laplacian

5.2.2.3 The Marr-Hildreth Operator

5.2.3 Other Handcrafted Edge Detection Operators

5.2.4 Comparison of Handcrafted Edge Detection Operators

5.2.5 Deep Learning Edge Detection Operators

5.2.6 Further Reading on Edge Detection

5.3 Phase Congruency

5.4 Localised Feature Extraction

5.4.1 Detecting Image Curvature (Corner Extraction)

5.4.1.1 Definition of Curvature

5.4.1.2 Measuring Curvature by Changes in Intensity (Differentiation)

5.4.1.3 Moravec and Harris Detectors

5.4.1.4 Further Reading on Curvature

5.4.2 Feature Point Detection; Region/Patch Analysis

5.4.2.1 Scale Invariant Feature Transform (SIFT)

5.4.2.2 Speeded Up Robust Features (SURF)

5.4.2.3 FAST, ORB, FREAK, LOCKY and other Handcrafted Keypoint Detectors

5.4.2.4 Deep Learning-Based Feature Detectors

5.4.3 Saliency

5.4.3.1 Basic Saliency

5.4.3.2 Context Aware Saliency

5.4.3.3 Other Handcrafted Saliency

Operators

5.4.3.4 Saliency by Deep Learning

5.5 Describing Image Motion

5.5.1 Area-based approach

5.5.2 Differential approach

5.5.3 Advancing Optical Flow: Deepflow, Epicflow and extensions; Acceleration

5.5.4 Analysis of Optical Flow

5.5.5 Optical Flow by Deep Learning

5.6 Further Reading

5.7 Chapter 5 References

6 High-Level Feature Extraction: Fixed Shape Analysis

6.1 Overview

6.2 Segmenting Objects from Images

6.2.1 Segmentation by Thresholding and Subtraction

6.2.2 Segmentation by Deep Learning

6.2.2.1 Semantic Segmentation by Convolutional Networks, with Extensions

6.2.2.2 Generic Segmentation

6.2.2.3 Matting

6.3 Feature Extraction by Low Level Features

6.3.1 Appearance-Based Approaches

6.3.1.1 Object Detection by Templates

6.3.1.2 Object Detection by Combinations of Parts

6.3.2 Distribution-Based Descriptors

6.3.2.1 Description by Interest Points (SIFT, SURF, BRIEF)

6.3.2.2 Characterising Object Appearance and Shape

- 6.3.3 Learning-Based Feature Descriptors
- 6.4 Feature Extraction by Shape
 - 6.4.1 Template Matching
 - 6.4.1.1 Definition
 - 6.4.1.2 Fourier Transform Implementation
 - 6.4.1.3 Discussion of Template Matching
 - 6.4.2 Shape Extraction/ Object Detection by Deep Learning
- 6.5 Hough Transform (HT)
 - 6.5.1 Overview
 - 6.5.2 HT for Lines
 - 6.5.3 HT for Circles
 - 6.5.4 HT for Ellipses
 - 6.5.5 Parameter Space Decomposition
 - 6.5.5.1 Parameter space reduction for lines
 - 6.5.5.1 Parameter space reduction for circles
 - 6.5.5.3 Parameter space reduction for ellipses
 - 6.5.6 Generalised Hough Transform (GHT)
 - 6.5.6.1 Formal Definition of the GHT
 - 6.5.6.2 Polar definition
 - 6.5.6.3 The GHT Technique
 - 6.5.6.4 Invariant GHT
 - 6.5.7 Other Extensions to the HT
 - 6.5.8 HT and Deep Learning
- 6.6 Further Reading
- 6.7 Chapter 6 References

7 High-Level Feature Extraction: Deformable Shape Analysis

- 7.1 Overview
- 7.2 Deformable Shape Analysis
 - 7.2.1 Deformable Templates
 - 7.2.2 Parts-based Shape Analysis
- 7.3 Active Contours (Snakes)
 - 7.3.1 Basics
 - 7.3.2 The Greedy Algorithm for Snakes

- 7.3.3 Complete (Kass) Snake Implementation
- 7.3.4 Other Snake Approaches
- 7.3.5 Further Snake Developments
- 7.3.6 Geometric Active Contours (Level-Set Based Approaches)
- 7.3.7 Deep Learning based Active Contours
- 7.4 Shape Skeletonisation
 - 7.4.1 Distance Transforms
 - 7.4.2 Symmetry
 - 7.4.3 Skeletonisation in the Deep Era
- 7.5 Flexible Shape Models – Active Shape and Active Appearance
- 7.6 Further Reading
- 7.7 Chapter 7 References

8 Object Description

- 8.1 Overview and Invariance Requirements
- 8.2 Boundary Descriptions
 - 8.2.1 Boundary and Region
 - 8.2.2 Chain Codes
 - 8.2.3 Fourier Descriptors
 - 8.2.3.1 Basis of Fourier Descriptors
 - 8.2.3.2 Fourier Expansion
 - 8.2.3.3 Shift invariance
 - 8.2.3.4 Discrete computation
 - 8.2.3.5 Cumulative Angular Function
 - 8.2.3.6 Elliptic Fourier Descriptors
 - 8.2.3.7 Invariance
- 8.3 Region Descriptors
 - 8.3.1 Basic Region Descriptors
 - 8.3.2 Moments
 - 8.3.2.1 Definition and Properties
 - 8.3.2.2 Geometric Moments
 - 8.3.2.3 Geometric Complex Moments and Centralised Moments
 - 8.3.2.4 Rotation and Scale Invariant Moments
 - 8.3.2.5 Zernike Moments
 - 8.3.2.6 Tchebichef Moments

- 8.3.2.7 Krawtchouk Moments
- 8.3.2.8 Other Moments
- 8.4 Further Reading
- 8.5 Chapter 8 References

9 Region Based Analysis

- 9.1 Overview
- 9.2 Advanced Denoising Operators and Image Super-Resolution
 - 9.2.1 Region-based Filtering
 - 9.2.1.1 Non-Local Means/ Self-Similarity
 - 9.2.1.2 Non-Local Means by Deep Learning
 - 9.2.1.3 Bilateral Filtering
 - 9.2.1.4 Anisotropic Diffusion
 - 9.2.1.5 Comparison of Handcrafted Smoothing Operators
 - 9.2.2 Image Denoising and Super-resolution by Deep Learning
 - 9.2.3 Other Processing Operators
 - 9.2.3.1 Force Field Transform
 - 9.2.3.2 Image Ray Transform
- 9.3 Region-Based Analysis
 - 9.3.1 Watershed transform
 - 9.3.2 Maximally Stable Extremal Regions
 - 9.3.3 Superpixels
 - 9.3.3.1 Basic Techniques and Normalised Cuts
 - 9.3.3.2 Simple Linear Iterative Clustering (SLIC)
 - 9.3.3.3 Superpixels by Deep Learning
- 9.4 Texture Description and Analysis
 - 9.4.1 What is Texture?
 - 9.4.2 Performance Requirements
 - 9.4.3 Structural Approaches
 - 9.4.4 Statistical Approaches
 - 9.4.4.1 Co-occurrence Matrix
 - 9.4.4.2 Machine learning based approaches
 - 9.4.5 Combination Approaches
 - 9.4.6 Local Binary Patterns
 - 9.4.7 Deep Learning for Texture
 - 9.4.8 Other Approaches

- 9.5 Further Reading
- 9.6 Chapter 9 References

10 Moving Object Detection, Description and Tracking

- 10.1 Overview
- 10.2 Moving Object Detection
 - 10.2.1 Basic Handcrafted Approaches
 - 10.2.1.1 Detection by Subtracting the Background
 - 10.2.1.2 Improving Quality by Morphology
 - 10.2.2 Modelling and Adapting to the (Static) Background
 - 10.2.3 Background Segmentation by Thresholding
 - 10.2.4 Problems and Advances: in comes Deep Learning
- 10.3 Tracking Moving Features
 - 10.3.1 Tracking Moving Objects
 - 10.3.2 Tracking by Local Search
 - 10.3.3 Problems in Tracking
 - 10.3.4 Approaches to Tracking
 - 10.3.5 MeanShift and Camshift
 - 10.3.5.1 Kernel-Based Density Estimation
 - 10.3.5.2 MeanShift Tracking
 - 10.3.5.3 Camshift Technique
 - 10.3.6 Deep tracking
- 10.4 Moving Feature Extraction and Description
 - 10.4.1 Moving (Biological) Shape Analysis
 - 10.4.2 Space-Time Interest Points
 - 10.4.3 Detecting Moving Shapes by Shape Matching in Image Sequences
 - 10.4.4 Moving Shape Description
 - 10.4.5 Deep Learning for Moving Shape Recognition
- 10.5 Further Reading
- 10.6 Chapter 10 References

11 Camera Geometry Fundamentals

- 11.1 Overview
- 11.2 Projective Space

11.2.1 Homogeneous co-ordinates and Projective Geometry	12.4.2 Transformation between RGB Models
11.2.2 Representation of a Line, Duality and Ideal Points	12.4.3 Transformation between RGB and CMY Models
11.2.3 Transformations in the Projective Space	12.5 Luminance and Chrominance Colour Models
11.2.4 Computing a Planar Homography	12.5.1 YUV, YIQ and YCbCr Models
11.3 The Perspective Camera	12.5.2 Luminance and Gamma Correction
11.3.1 Perspective Camera Model	12.5.3 Chrominance
11.3.2 Parameters of the Perspective Camera Model	12.5.4 Transformations between YUV, YIQ and RGB Colour Models
11.3.3 Computing a Projection from an Image	12.5.5 Colour Model for Component Video: YPbPr
11.4 Affine Camera	12.5.6 Colour Model for Digital Video: YCbCr
11.4.1 Affine Camera Model	12.6 Additive Perceptual Colour Models
11.4.2 Affine Camera Model and the Perspective Projection	12.6.1 The HSV and HLS Colour Models
11.4.3 Parameters of the Affine Camera Model	12.6.2 The Hexagonal Model: HSV
11.5 Weak Perspective Model	12.6.3 The Triangular Model: HLS
11.6 Discussion	12.6.4 Transformation between HLS and RGB
11.7 Further Reading	12.7 More Colour Models
11.8 Chapter 11 References	12.8 Chapter 12 References
12 Colour Images	
12.1 Overview	
12.2 Colour Image Theory	
12.2.1 Colour Images	
12.2.2 Tristimulus Theory	
12.2.3 The Colourimetric Equation	
12.2.4 Luminosity Function	
12.3 Perception based Colour Models: CIE RGB and CIE XYZ	
12.3.1 CIE RGB Colour Model: Wright-Guild Data	
12.3.2 CIE RGB Colour Matching Functions	
12.3.3 CIE RGB Chromaticity Diagram and Chromaticity Coordinates	
12.3.4 CIE XYZ Colour Model	
12.3.5 CIE XYZ Colour Matching Functions	
12.3.6 XYZ Chromaticity Diagram	
12.4 Additive and Subtractive Colour Models	
12.4.1 RGB and CMY	

Index