

<<计算机视觉>>

图书基本信息

书名：<<计算机视觉>>

13位ISBN编号：9787121168307

10位ISBN编号：7121168308

出版时间：2012-5

出版时间：电子工业出版社

作者：福赛斯(David A. Forsyth),泊斯(Jean Ponce)

页数：761

字数：1268000

译者：David A.Forsyth

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

<<计算机视觉>>

内容概要

计算机视觉是研究如何使人工系统从图像或多维数据中“感知”的科学。本书是计算机视觉领域的经典教材，内容涉及几何摄像模型、光照和着色、色彩、线性滤波、局部图像特征、纹理、立体相对、运动结构、聚类分割、组合与模型拟合、追踪、配准、平滑表面与骨架、距离数据、图像分类、对象检测与识别、基于图像的建模与渲染、人形研究、图像搜索与检索、优化技术等内容。

与前一版相比，本书简化了部分主题，增加了应用示例，重写了关于现代特性的内容，详述了现代图像编辑技术与对象识别技术。

<<计算机视觉>>

作者简介

作者：（美国）福赛斯（David A. Forsyth）（美国）泊斯（Jean Ponce）

书籍目录

I IMAGE FORMATION

1 Geometric Camera Models

1.1 Image Formation

1.1.1 Pinhole Perspective

1.1.2 Weak Perspective

1.1.3 Cameras with Lenses

1.1.4 The Human Eye

1.2 Intrinsic and Extrinsic Parameters

1.2.1 Rigid Transformations and Homogeneous Coordinates

1.2.2 Intrinsic Parameters

1.2.3 Extrinsic Parameters

1.2.4 Perspective Projection Matrices

1.2.5 Weak-Perspective Projection Matrices

1.3 Geometric Camera Calibration

1.3.1 A Linear Approach to Camera Calibration

1.3.2 A Nonlinear Approach to Camera Calibration

1.4 Notes

2 Light and Shading

2.1 Modelling Pixel Brightness

2.1.1 Reflection at Surfaces

2.1.2 Sources and Their Effects

2.1.3 The Lambertian+Specular Model

2.1.4 Area Sources

2.2 Inference from Shading

2.2.1 Radiometric Calibration and High Dynamic Range Images

2.2.2 The Shape of Specularities

2.2.3 Inferring Lightness and Illumination

2.2.4 Photometric Stereo: Shape from Multiple Shaded Images

2.3 Modelling Interreflection

2.3.1 The Illumination at a Patch Due to an Area Source

2.3.2 Radiosity and Exitance

2.3.3 An Interreflection Model

2.3.4 Qualitative Properties of Interreflections

2.4 Shape from One Shaded Image

2.5 Notes

3 Color

3.1 Human Color Perception

3.1.1 Color Matching

3.1.2 Color Receptors

3.2 The Physics of Color

3.2.1 The Color of Light Sources

3.2.2 The Color of Surfaces

3.3 Representing Color

3.3.1 Linear Color Spaces

3.3.2 Non-linear Color Spaces

<<计算机视觉>>

3.4 A Model of Image Color

3.4.1 The Diffuse Term

3.4.2 The Specular Term

3.5 Inference from Color

3.5.1 Finding Specularities Using Color

3.5.2 Shadow Removal Using Color

3.5.3 Color Constancy: Surface Color from Image Color

3.6 Notes

II EARLY VISION: JUST ONE IMAGE

4 Linear Filters

4.1 Linear Filters and Convolution

4.1.1 Convolution

4.2 Shift Invariant Linear Systems

4.2.1 Discrete Convolution

4.2.2 Continuous Convolution

4.2.3 Edge Effects in Discrete Convolutions

4.3 Spatial Frequency and Fourier Transforms

4.3.1 Fourier Transforms

4.4 Sampling and Aliasing

4.4.1 Sampling

4.4.2 Aliasing

4.4.3 Smoothing and Resampling

4.5 Filters as Templates

4.5.1 Convolution as a Dot Product

4.5.2 Changing Basis

4.6 Technique: Normalized Correlation and Finding Patterns

4.6.1 Controlling the Television by Finding Hands by

Normalized

Correlation

4.7 Technique: Scale and Image Pyramids

4.7.1 The Gaussian Pyramid

4.7.2 Applications of Scaled Representations

4.8 Notes

5 Local Image Features

5.1 Computing the Image Gradient

5.1.1 Derivative of Gaussian Filters

5.2 Representing the Image Gradient

5.2.1 Gradient-Based Edge Detectors

5.2.2 Orientations

5.3 Finding Corners and Building Neighborhoods

5.3.1 Finding Corners

5.3.2 Using Scale and Orientation to Build a Neighborhood

5.4 Describing Neighborhoods with SIFT and HOG Features

5.4.1 SIFT Features

5.4.2 HOG Features

5.5 Computing Local Features in Practice

5.6 Notes

<<计算机视觉>>

6 Texture

- 6.1 Local Texture Representations Using Filters
 - 6.1.1 Spots and Bars
 - 6.1.2 From Filter Outputs to Texture Representation
 - 6.1.3 Local Texture Representations in Practice
- 6.2 Pooled Texture Representations by Discovering Textons
 - 6.2.1 Vector Quantization and Textons
 - 6.2.2 K-means Clustering for Vector Quantization
- 6.3 Synthesizing Textures and Filling Holes in Images
 - 6.3.1 Synthesis by Sampling Local Models
 - 6.3.2 Filling in Holes in Images
- 6.4 Image Denoising
 - 6.4.1 Non-local Means
 - 6.4.2 Block Matching 3D (BM3D)
 - 6.4.3 Learned Sparse Coding
 - 6.4.4 Results
- 6.5 Shape from Texture
 - 6.5.1 Shape from Texture for Planes
 - 6.5.2 Shape from Texture for Curved Surfaces
- 6.6 Notes

III EARLY VISION: MULTIPLE IMAGES

7 Stereopsis

- 7.1 Binocular Camera Geometry and the Epipolar Constraint
 - 7.1.1 Epipolar Geometry
 - 7.1.2 The Essential Matrix
 - 7.1.3 The Fundamental Matrix
- 7.2 Binocular Reconstruction
 - 7.2.1 Image Rectification
- 7.3 Human Stereopsis
- 7.4 Local Methods for Binocular Fusion
 - 7.4.1 Correlation
 - 7.4.2 Multi-Scale Edge Matching
- 7.5 Global Methods for Binocular Fusion
 - 7.5.1 Ordering Constraints and Dynamic Programming
 - 7.5.2 Smoothness and Graphs
- 7.6 Using More Cameras
 - 7.7 Application: Robot Navigation
- 7.8 Notes

8 Structure from Motion

- 8.1 Internally Calibrated Perspective Cameras
 - 8.1.1 Natural Ambiguity of the Problem
 - 8.1.2 Euclidean Structure and Motion from Two Images
 - 8.1.3 Euclidean Structure and Motion from Multiple Images
- 8.2 Uncalibrated Weak-Perspective Cameras
 - 8.2.1 Natural Ambiguity of the Problem
 - 8.2.2 Affine Structure and Motion from Two Images
 - 8.2.3 Affine Structure and Motion from Multiple Images

<<计算机视觉>>

- 8.2.4 From Affine to Euclidean Shape
- 8.3 Uncalibrated Perspective Cameras
 - 8.3.1 Natural Ambiguity of the Problem
 - 8.3.2 Projective Structure and Motion from Two Images
 - 8.3.3 Projective Structure and Motion from Multiple Images
 - 8.3.4 From Projective to Euclidean Shape
- 8.4 Notes
- IV MID-LEVEL VISION
- 9 Segmentation by Clustering
 - 9.1 Human Vision: Grouping and Gestalt
 - 9.2 Important Applications
 - 9.2.1 Background Subtraction
 - 9.2.2 Shot Boundary Detection
 - 9.2.3 Interactive Segmentation
 - 9.2.4 Forming Image Regions
 - 9.3 Image Segmentation by Clustering Pixels
 - 9.3.1 Basic Clustering Methods
 - 9.3.2 The Watershed Algorithm
 - 9.3.3 Segmentation Using K-means
 - 9.3.4 Mean Shift: Finding Local Modes in Data
 - 9.3.5 Clustering and Segmentation with Mean Shift
 - 9.4 Segmentation, Clustering, and Graphs
 - 9.4.1 Terminology and Facts for Graphs
 - 9.4.2 Agglomerative Clustering with a Graph
 - 9.4.3 Divisive Clustering with a Graph
 - 9.4.4 Normalized Cuts
 - 9.5 Image Segmentation in Practice
 - 9.5.1 Evaluating Segmenters
 - 9.6 Notes
- 10 Grouping and Model Fitting
 - 10.1 The Hough Transform
 - 10.1.1 Fitting Lines with the Hough Transform
 - 10.1.2 Using the Hough Transform
 - 10.2 Fitting Lines and Planes
 - 10.2.1 Fitting a Single Line
 - 10.2.2 Fitting Planes
 - 10.2.3 Fitting Multiple Lines
 - 10.3 Fitting Curved Structures
 - 10.4 Robustness
 - 10.4.1 M-Estimators
 - 10.4.2 RANSAC: Searching for Good Points
 - 10.5 Fitting Using Probabilistic Models
 - 10.5.1 Missing Data Problems
 - 10.5.2 Mixture Models and Hidden Variables
 - 10.5.3 The EM Algorithm for Mixture Models
 - 10.5.4 Difficulties with the EM Algorithm
 - 10.6 Motion Segmentation by Parameter Estimation

<<计算机视觉>>

- 10.6.1 Optical Flow and Motion
- 10.6.2 Flow Models
- 10.6.3 Motion Segmentation with Layers
- 10.7 Model Selection: Which Model Is the Best Fit?
 - 10.7.1 Model Selection Using Cross-Validation
- 10.8 Notes
- 11 Tracking
 - 11.1 Simple Tracking Strategies
 - 11.1.1 Tracking by Detection
 - 11.1.2 Tracking Translations by Matching
 - 11.1.3 Using Affine Transformations to Confirm a Match
 - 11.2 Tracking Using Matching
 - 11.2.1 Matching Summary Representations
 - 11.2.2 Tracking Using Flow
 - 11.3 Tracking Linear Dynamical Models with Kalman Filters
 - 11.3.1 Linear Measurements and Linear Dynamics
 - 11.3.2 The Kalman Filter
 - 11.3.3 Forward-backward Smoothing
 - 11.4 Data Association
 - 11.4.1 Linking Kalman Filters with Detection Methods
 - 11.4.2 Key Methods of Data Association
 - 11.5 Particle Filtering
 - 11.5.1 Sampled Representations of Probability Distributions
 - 11.5.2 The Simplest Particle Filter
 - 11.5.3 The Tracking Algorithm
 - 11.5.4 A Workable Particle Filter
 - 11.5.5 Practical Issues in Particle Filters
 - 11.6 Notes
- V HIGH-LEVEL VISION
- 12 Registration
 - 12.1 Registering Rigid Objects
 - 12.1.1 Iterated Closest Points
 - 12.1.2 Searching for Transformations via Correspondences
 - 12.1.3 Application: Building Image Mosaics
 - 12.2 Model-based Vision: Registering Rigid Objects with Projection
 - 12.2.1 Verification: Comparing Transformed and Rendered Source to Target
 - 12.3 Registering Deformable Objects
 - 12.3.1 Deforming Texture with Active Appearance Models
 - 12.3.2 Active Appearance Models in Practice
 - 12.3.3 Application: Registration in Medical Imaging Systems
 - 12.4 Notes
- 13 Smooth Surfaces and Their Outlines
 - 13.1 Elements of Differential Geometry
 - 13.1.1 Curves

<<计算机视觉>>

- 13.1.2 Surfaces
- 13.2 Contour Geometry
 - 13.2.1 The Occluding Contour and the Image Contour
 - 13.2.2 The Cusps and Inflections of the Image Contour
 - 13.2.3 Koenderink ' s Theorem
- 13.3 Visual Events: More Differential Geometry
 - 13.3.1 The Geometry of the Gauss Map
 - 13.3.2 Asymptotic Curves
 - 13.3.3 The Asymptotic Spherical Map
 - 13.3.4 Local Visual Events
 - 13.3.5 The Bitangent Ray Manifold
 - 13.3.6 Multilocal Visual Events
 - 13.3.7 The Aspect Graph
- 13.4 Notes
- 14 Range Data
 - 14.1 Active Range Sensors
 - 14.2 Range Data Segmentation
 - 14.2.1 Elements of Analytical Differential Geometry
 - 14.2.2 Finding Step and Roof Edges in Range Images
 - 14.2.3 Segmenting Range Images into Planar Regions
 - 14.3 Range Image Registration and Model Acquisition
 - 14.3.1 Quaternions
 - 14.3.2 Registering Range Images
 - 14.3.3 Fusing Multiple Range Images
 - 14.4 Object Recognition
 - 14.4.1 Matching Using Interpretation Trees
 - 14.4.2 Matching Free-Form Surfaces Using Spin Images
 - 14.5 Kinect
 - 14.5.1 Features
 - 14.5.2 Technique: Decision Trees and Random Forests
 - 14.5.3 Labeling Pixels
 - 14.5.4 Computing Joint Positions
 - 14.6 Notes
- 15 Learning to Classify
 - 15.1 Classification, Error, and Loss
 - 15.1.1 Using Loss to Determine Decisions
 - 15.1.2 Training Error, Test Error, and Overfitting
 - 15.1.3 Regularization
 - 15.1.4 Error Rate and Cross-Validation
 - 15.1.5 Receiver Operating Curves
 - 15.2 Major Classification Strategies
 - 15.2.1 Example: Mahalanobis Distance
 - 15.2.2 Example: Class-Conditional Histograms and Naive Bayes
 - 15.2.3 Example: Classification Using Nearest Neighbors
 - 15.2.4 Example: The Linear Support Vector Machine
 - 15.2.5 Example: Kernel Machines

<<计算机视觉>>

- 15.2.6 Example: Boosting and Adaboost
- 15.3 Practical Methods for Building Classifiers
 - 15.3.1 Manipulating Training Data to Improve Performance
 - 15.3.2 Building Multi-Class Classifiers Out of Binary Classifiers
 - 15.3.3 Solving for SVMs and Kernel Machines
- 15.4 Notes
- 16 Classifying Images
 - 16.1 Building Good Image Features
 - 16.1.1 Example Applications
 - 16.1.2 Encoding Layout with GIST Features
 - 16.1.3 Summarizing Images with Visual Words
 - 16.1.4 The Spatial Pyramid Kernel
 - 16.1.5 Dimension Reduction with Principal Components
 - 16.1.6 Dimension Reduction with Canonical Variates
 - 16.1.7 Example Application: Identifying Explicit Images
 - 16.1.8 Example Application: Classifying Materials
 - 16.1.9 Example Application: Classifying Scenes
 - 16.2 Classifying Images of Single Objects
 - 16.2.1 Image Classification Strategies
 - 16.2.2 Evaluating Image Classification Systems
 - 16.2.3 Fixed Sets of Classes
 - 16.2.4 Large Numbers of Classes
 - 16.2.5 Flowers, Leaves, and Birds: Some Specialized
- Problems
 - 16.3 Image Classification in Practice
 - 16.3.1 Codes for Image Features
 - 16.3.2 Image Classification Datasets
 - 16.3.3 Dataset Bias
 - 16.3.4 Crowdsourcing Dataset Collection
 - 16.4 Notes
- 17 Detecting Objects in Images
 - 17.1 The Sliding Window Method
 - 17.1.1 Face Detection
 - 17.1.2 Detecting Humans
 - 17.1.3 Detecting Boundaries
 - 17.2 Detecting Deformable Objects
 - 17.3 The State of the Art of Object Detection
 - 17.3.1 Datasets and Resources
 - 17.4 Notes
- 18 Topics in Object Recognition
 - 18.1 What Should Object Recognition Do?
 - 18.1.1 What Should an Object Recognition System Do?
 - 18.1.2 Current Strategies for Object Recognition
 - 18.1.3 What Is Categorization?
 - 18.1.4 Selection: What Should Be Described?
 - 18.2 Feature Questions

<<计算机视觉>>

- 18.2.1 Improving Current Image Features
- 18.2.2 Other Kinds of Image Feature
- 18.3 Geometric Questions
- 18.4 Semantic Questions
 - 18.4.1 Attributes and the Unfamiliar
 - 18.4.2 Parts, Poselets and Consistency
 - 18.4.3 Chunks of Meaning
- VI APPLICATIONS AND TOPICS
- 19 Image-Based Modeling and Rendering
 - 19.1 Visual Hulls
 - 19.1.1 Main Elements of the Visual Hull Model
 - 19.1.2 Tracing Intersection Curves
 - 19.1.3 Clipping Intersection Curves
 - 19.1.4 Triangulating Cone Strips
 - 19.1.5 Results
 - 19.1.6 Going Further: Carved Visual Hulls
 - 19.2 Patch-Based Multi-View Stereopsis
 - 19.2.1 Main Elements of the PMVS Model
 - 19.2.2 Initial Feature Matching
 - 19.2.3 Expansion
 - 19.2.4 Filtering
 - 19.2.5 Results
 - 19.3 The Light Field
 - 19.4 Notes
- 20 Looking at People
 - 20.1 HMM's, Dynamic Programming, and Tree-Structured Models
 - 20.1.1 Hidden Markov Models
 - 20.1.2 Inference for an HMM
 - 20.1.3 Fitting an HMM with EM
 - 20.1.4 Tree-Structured Energy Models
 - 20.2 Parsing People in Images
 - 20.2.1 Parsing with Pictorial Structure Models
 - 20.2.2 Estimating the Appearance of Clothing
 - 20.3 Tracking People
 - 20.3.1 Why Human Tracking Is Hard
 - 20.3.2 Kinematic Tracking by Appearance
 - 20.3.3 Kinematic Human Tracking Using Templates
 - 20.4 3D from 2D: Lifting
 - 20.4.1 Reconstruction in an Orthographic View
 - 20.4.2 Exploiting Appearance for Unambiguous Reconstructions
 - 20.4.3 Exploiting Motion for Unambiguous Reconstructions
 - 20.5 Activity Recognition
 - 20.5.1 Background: Human Motion Data
 - 20.5.2 Body Configuration and Activity Recognition
 - 20.5.3 Recognizing Human Activities with Appearance
- Features

<<计算机视觉>>

- 20.5.4 Recognizing Human Activities with Compositional Models
- 20.6 Resources
- 20.7 Notes
- 21 Image Search and Retrieval
 - 21.1 The Application Context
 - 21.1.1 Applications
 - 21.1.2 User Needs
 - 21.1.3 Types of Image Query
 - 21.1.4 What Users Do with Image Collections
 - 21.2 Basic Technologies from Information Retrieval
 - 21.2.1 Word Counts
 - 21.2.2 Smoothing Word Counts
 - 21.2.3 Approximate Nearest Neighbors and Hashing
 - 21.2.4 Ranking Documents
 - 21.3 Images as Documents
 - 21.3.1 Matching Without Quantization
 - 21.3.2 Ranking Image Search Results
 - 21.3.3 Browsing and Layout
 - 21.3.4 Laying Out Images for Browsing
 - 21.4 Predicting Annotations for Pictures
 - 21.4.1 Annotations from Nearby Words
 - 21.4.2 Annotations from the Whole Image
 - 21.4.3 Predicting Correlated Words with Classifiers
 - 21.4.4 Names and Faces
 - 21.4.5 Generating Tags with Segments
 - 21.5 The State of the Art of Word Prediction
 - 21.5.1 Resources
 - 21.5.2 Comparing Methods
 - 21.5.3 Open Problems
 - 21.6 Notes
- VII BACKGROUND MATERIAL
- 22 Optimization Techniques
 - 22.1 Linear Least-Squares Methods
 - 22.1.1 Normal Equations and the Pseudoinverse
 - 22.1.2 Homogeneous Systems and Eigenvalue Problems
 - 22.1.3 Generalized Eigenvalues Problems
 - 22.1.4 An Example: Fitting a Line to Points in a Plane
 - 22.1.5 Singular Value Decomposition
 - 22.2 Nonlinear Least-Squares Methods
 - 22.2.1 Newton ' s Method: Square Systems of Nonlinear Equations.
 - 22.2.2 Newton ' s Method for Overconstrained Systems
 - 22.2.3 The Gauss—Newton and Levenberg—Marquardt Algorithms
 - 22.3 Sparse Coding and Dictionary Learning
 - 22.3.1 Sparse Coding
 - 22.3.2 Dictionary Learning

<<计算机视觉>>

22.3.3 Supervised Dictionary Learning

22.4 Min-Cut/Max-Flow Problems and Combinatorial Optimization

22.4.1 Min-Cut Problems

22.4.2 Quadratic Pseudo-Boolean Functions

22.4.3 Generalization to Integer Variables

22.5 Notes

Bibliography

Index

List of Algorithms

Courses

Computer Vision (Computer Science)

Previous Edition(s)

Net price is Pearson ' s wholesale price to college bookstores and other resellers.

Table of Contents

I IMAGE FORMATION

1 Geometric Camera Models

1.1 Image Formation

1.1.1 Pinhole Perspective

1.1.2 Weak Perspective

1.1.3 Cameras with Lenses

1.1.4 The Human Eye

1.2 Intrinsic and Extrinsic Parameters

1.2.1 Rigid Transformations and Homogeneous Coordinates

1.2.2 Intrinsic Parameters

1.2.3 Extrinsic Parameters

1.2.4 Perspective Projection Matrices

1.2.5 Weak-Perspective Projection Matrices

1.3 Geometric Camera Calibration

1.3.1 A Linear Approach to Camera Calibration

1.3.2 A Nonlinear Approach to Camera Calibration

1.4 Notes

2 Light and Shading

2.1 Modelling Pixel Brightness

2.1.1 Reflection at Surfaces

2.1.2 Sources and Their Effects

2.1.3 The Lambertian+Specular Model

2.1.4 Area Sources

2.2 Inference from Shading

2.2.1 Radiometric Calibration and High Dynamic Range Images

2.2.2 The Shape of Specularities

2.2.3 Inferring Lightness and Illumination

2.2.4 Photometric Stereo: Shape from Multiple Shaded Images

2.3 Modelling Interreflection

2.3.1 The Illumination at a Patch Due to an Area Source

2.3.2 Radiosity and Exitance

<<计算机视觉>>

2.3.3 An Interreflection Model

2.3.4 Qualitative Properties of Interreflections

2.4 Shape from One Shaded Image

2.5 Notes

3 Color

3.1 Human Color Perception

3.1.1 Color Matching

3.1.2 Color Receptors

3.2 The Physics of Color

3.2.1 The Color of Light Sources

3.2.2 The Color of Surfaces

3.3 Representing Color

3.3.1 Linear Color Spaces

3.3.2 Non-linear Color Spaces

3.4 A Model of Image Color

3.4.1 The Diffuse Term

3.4.2 The Specular Term

3.5 Inference from Color

3.5.1 Finding Specularities Using Color

3.5.2 Shadow Removal Using Color

3.5.3 Color Constancy: Surface Color from Image Color

3.6 Notes

II EARLY VISION: JUST ONE IMAGE

4 Linear Filters

4.1 Linear Filters and Convolution

4.1.1 Convolution

4.2 Shift Invariant Linear Systems

4.2.1 Discrete Convolution

4.2.2 Continuous Convolution

4.2.3 Edge Effects in Discrete Convolutions

4.3 Spatial Frequency and Fourier Transforms

4.3.1 Fourier Transforms

4.4 Sampling and Aliasing

4.4.1 Sampling

4.4.2 Aliasing

4.4.3 Smoothing and Resampling

4.5 Filters as Templates

4.5.1 Convolution as a Dot Product

4.5.2 Changing Basis

4.6 Technique: Normalized Correlation and Finding Patterns

4.6.1 Controlling the Television by Finding Hands by

Normalized

Correlation

4.7 Technique: Scale and Image Pyramids

4.7.1 The Gaussian Pyramid

4.7.2 Applications of Scaled Representations

4.8 Notes

<<计算机视觉>>

5 Local Image Features

5.1 Computing the Image Gradient

5.1.1 Derivative of Gaussian Filters

5.2 Representing the Image Gradient

5.2.1 Gradient-Based Edge Detectors

5.2.2 Orientations

5.3 Finding Corners and Building Neighborhoods

5.3.1 Finding Corners

5.3.2 Using Scale and Orientation to Build a Neighborhood

5.4 Describing Neighborhoods with SIFT and HOG Features

5.4.1 SIFT Features

5.4.2 HOG Features

5.5 Computing Local Features in Practice

5.6 Notes

6 Texture

6.1 Local Texture Representations Using Filters

6.1.1 Spots and Bars

6.1.2 From Filter Outputs to Texture Representation

6.1.3 Local Texture Representations in Practice

6.2 Pooled Texture Representations by Discovering Textons

6.2.1 Vector Quantization and Textons

6.2.2 K-means Clustering for Vector Quantization

6.3 Synthesizing Textures and Filling Holes in Images

6.3.1 Synthesis by Sampling Local Models

6.3.2 Filling in Holes in Images

6.4 Image Denoising

6.4.1 Non-local Means

6.4.2 Block Matching 3D (BM3D)

6.4.3 Learned Sparse Coding

6.4.4 Results

6.5 Shape from Texture

6.5.1 Shape from Texture for Planes

6.5.2 Shape from Texture for Curved Surfaces

6.6 Notes

III EARLY VISION: MULTIPLE IMAGES

7 Stereopsis

7.1 Binocular Camera Geometry and the Epipolar Constraint

7.1.1 Epipolar Geometry

7.1.2 The Essential Matrix

7.1.3 The Fundamental Matrix

7.2 Binocular Reconstruction

7.2.1 Image Rectification

7.3 Human Stereopsis

7.4 Local Methods for Binocular Fusion

7.4.1 Correlation

7.4.2 Multi-Scale Edge Matching

7.5 Global Methods for Binocular Fusion

<<计算机视觉>>

- 7.5.1 Ordering Constraints and Dynamic Programming
- 7.5.2 Smoothness and Graphs
- 7.6 Using More Cameras
- 7.7 Application: Robot Navigation
- 7.8 Notes
- 8 Structure from Motion
 - 8.1 Internally Calibrated Perspective Cameras
 - 8.1.1 Natural Ambiguity of the Problem
 - 8.1.2 Euclidean Structure and Motion from Two Images
 - 8.1.3 Euclidean Structure and Motion from Multiple Images
 - 8.2 Uncalibrated Weak-Perspective Cameras
 - 8.2.1 Natural Ambiguity of the Problem
 - 8.2.2 Affine Structure and Motion from Two Images
 - 8.2.3 Affine Structure and Motion from Multiple Images
 - 8.2.4 From Affine to Euclidean Shape
 - 8.3 Uncalibrated Perspective Cameras
 - 8.3.1 Natural Ambiguity of the Problem
 - 8.3.2 Projective Structure and Motion from Two Images
 - 8.3.3 Projective Structure and Motion from Multiple Images
 - 8.3.4 From Projective to Euclidean Shape
 - 8.4 Notes
- IV MID-LEVEL VISION
- 9 Segmentation by Clustering
 - 9.1 Human Vision: Grouping and Gestalt
 - 9.2 Important Applications
 - 9.2.1 Background Subtraction
 - 9.2.2 Shot Boundary Detection
 - 9.2.3 Interactive Segmentation
 - 9.2.4 Forming Image Regions
 - 9.3 Image Segmentation by Clustering Pixels
 - 9.3.1 Basic Clustering Methods
 - 9.3.2 The Watershed Algorithm
 - 9.3.3 Segmentation Using K-means
 - 9.3.4 Mean Shift: Finding Local Modes in Data
 - 9.3.5 Clustering and Segmentation with Mean Shift
 - 9.4 Segmentation, Clustering, and Graphs
 - 9.4.1 Terminology and Facts for Graphs
 - 9.4.2 Agglomerative Clustering with a Graph
 - 9.4.3 Divisive Clustering with a Graph
 - 9.4.4 Normalized Cuts
 - 9.5 Image Segmentation in Practice
 - 9.5.1 Evaluating Segmenters
 - 9.6 Notes
- 10 Grouping and Model Fitting
 - 10.1 The Hough Transform
 - 10.1.1 Fitting Lines with the Hough Transform
 - 10.1.2 Using the Hough Transform

<<计算机视觉>>

- 10.2 Fitting Lines and Planes
 - 10.2.1 Fitting a Single Line
 - 10.2.2 Fitting Planes
 - 10.2.3 Fitting Multiple Lines
- 10.3 Fitting Curved Structures
- 10.4 Robustness
 - 10.4.1 M-Estimators
 - 10.4.2 RANSAC: Searching for Good Points
- 10.5 Fitting Using Probabilistic Models
 - 10.5.1 Missing Data Problems
 - 10.5.2 Mixture Models and Hidden Variables
 - 10.5.3 The EM Algorithm for Mixture Models
 - 10.5.4 Difficulties with the EM Algorithm
- 10.6 Motion Segmentation by Parameter Estimation
 - 10.6.1 Optical Flow and Motion
 - 10.6.2 Flow Models
 - 10.6.3 Motion Segmentation with Layers
- 10.7 Model Selection: Which Model Is the Best Fit?
 - 10.7.1 Model Selection Using Cross-Validation
- 10.8 Notes
- 11 Tracking
 - 11.1 Simple Tracking Strategies
 - 11.1.1 Tracking by Detection
 - 11.1.2 Tracking Translations by Matching
 - 11.1.3 Using Affine Transformations to Confirm a Match
 - 11.2 Tracking Using Matching
 - 11.2.1 Matching Summary Representations
 - 11.2.2 Tracking Using Flow
 - 11.3 Tracking Linear Dynamical Models with Kalman Filters
 - 11.3.1 Linear Measurements and Linear Dynamics
 - 11.3.2 The Kalman Filter
 - 11.3.3 Forward-backward Smoothing
 - 11.4 Data Association
 - 11.4.1 Linking Kalman Filters with Detection Methods
 - 11.4.2 Key Methods of Data Association
 - 11.5 Particle Filtering
 - 11.5.1 Sampled Representations of Probability Distributions
 - 11.5.2 The Simplest Particle Filter
 - 11.5.3 The Tracking Algorithm
 - 11.5.4 A Workable Particle Filter
 - 11.5.5 Practical Issues in Particle Filters
 - 11.6 Notes
- V HIGH-LEVEL VISION
- 12 Registration
 - 12.1 Registering Rigid Objects
 - 12.1.1 Iterated Closest Points
 - 12.1.2 Searching for Transformations via Correspondences

<<计算机视觉>>

- 12.1.3 Application: Building Image Mosaics
- 12.2 Model-based Vision: Registering Rigid Objects with Projection
 - 12.2.1 Verification: Comparing Transformed and Rendered Source to Target
- 12.3 Registering Deformable Objects
 - 12.3.1 Deforming Texture with Active Appearance Models
 - 12.3.2 Active Appearance Models in Practice
 - 12.3.3 Application: Registration in Medical Imaging Systems
- 12.4 Notes
- 13 Smooth Surfaces and Their Outlines
 - 13.1 Elements of Differential Geometry
 - 13.1.1 Curves
 - 13.1.2 Surfaces
 - 13.2 Contour Geometry
 - 13.2.1 The Occluding Contour and the Image Contour
 - 13.2.2 The Cusps and Inflections of the Image Contour
 - 13.2.3 Koenderink ' s Theorem
 - 13.3 Visual Events: More Differential Geometry
 - 13.3.1 The Geometry of the Gauss Map
 - 13.3.2 Asymptotic Curves
 - 13.3.3 The Asymptotic Spherical Map
 - 13.3.4 Local Visual Events
 - 13.3.5 The Bitangent Ray Manifold
 - 13.3.6 Multilocal Visual Events
 - 13.3.7 The Aspect Graph
 - 13.4 Notes
- 14 Range Data
 - 14.1 Active Range Sensors
 - 14.2 Range Data Segmentation
 - 14.2.1 Elements of Analytical Differential Geometry
 - 14.2.2 Finding Step and Roof Edges in Range Images
 - 14.2.3 Segmenting Range Images into Planar Regions
 - 14.3 Range Image Registration and Model Acquisition
 - 14.3.1 Quaternions
 - 14.3.2 Registering Range Images
 - 14.3.3 Fusing Multiple Range Images
 - 14.4 Object Recognition
 - 14.4.1 Matching Using Interpretation Trees
 - 14.4.2 Matching Free-Form Surfaces Using Spin Images
 - 14.5 Kinect
 - 14.5.1 Features
 - 14.5.2 Technique: Decision Trees and Random Forests
 - 14.5.3 Labeling Pixels
 - 14.5.4 Computing Joint Positions
 - 14.6 Notes

<<计算机视觉>>

15 Learning to Classify

15.1 Classification, Error, and Loss

- 15.1.1 Using Loss to Determine Decisions
- 15.1.2 Training Error, Test Error, and Overfitting
- 15.1.3 Regularization
- 15.1.4 Error Rate and Cross-Validation
- 15.1.5 Receiver Operating Curves

15.2 Major Classification Strategies

- 15.2.1 Example: Mahalanobis Distance
- 15.2.2 Example: Class-Conditional Histograms and Naive

Bayes

- 15.2.3 Example: Classification Using Nearest Neighbors
- 15.2.4 Example: The Linear Support Vector Machine
- 15.2.5 Example: Kernel Machines
- 15.2.6 Example: Boosting and Adaboost

15.3 Practical Methods for Building Classifiers

- 15.3.1 Manipulating Training Data to Improve Performance
- 15.3.2 Building Multi-Class Classifiers Out of Binary

Classifiers

- 15.3.3 Solving for SVMs and Kernel Machines

15.4 Notes

16 Classifying Images

16.1 Building Good Image Features

- 16.1.1 Example Applications
- 16.1.2 Encoding Layout with GIST Features
- 16.1.3 Summarizing Images with Visual Words
- 16.1.4 The Spatial Pyramid Kernel
- 16.1.5 Dimension Reduction with Principal Components
- 16.1.6 Dimension Reduction with Canonical Variates
- 16.1.7 Example Application: Identifying Explicit Images
- 16.1.8 Example Application: Classifying Materials
- 16.1.9 Example Application: Classifying Scenes

16.2 Classifying Images of Single Objects

- 16.2.1 Image Classification Strategies
- 16.2.2 Evaluating Image Classification Systems
- 16.2.3 Fixed Sets of Classes
- 16.2.4 Large Numbers of Classes
- 16.2.5 Flowers, Leaves, and Birds: Some Specialized

Problems

16.3 Image Classification in Practice

- 16.3.1 Codes for Image Features
- 16.3.2 Image Classification Datasets
- 16.3.3 Dataset Bias
- 16.3.4 Crowdsourcing Dataset Collection

16.4 Notes

17 Detecting Objects in Images

17.1 The Sliding Window Method

<<计算机视觉>>

- 17.1.1 Face Detection
- 17.1.2 Detecting Humans
- 17.1.3 Detecting Boundaries
- 17.2 Detecting Deformable Objects
- 17.3 The State of the Art of Object Detection
 - 17.3.1 Datasets and Resources
- 17.4 Notes
- 18 Topics in Object Recognition
 - 18.1 What Should Object Recognition Do?
 - 18.1.1 What Should an Object Recognition System Do?
 - 18.1.2 Current Strategies for Object Recognition
 - 18.1.3 What Is Categorization?
 - 18.1.4 Selection: What Should Be Described?
 - 18.2 Feature Questions
 - 18.2.1 Improving Current Image Features
 - 18.2.2 Other Kinds of Image Feature
 - 18.3 Geometric Questions
 - 18.4 Semantic Questions
 - 18.4.1 Attributes and the Unfamiliar
 - 18.4.2 Parts, Poselets and Consistency
 - 18.4.3 Chunks of Meaning
- VI APPLICATIONS AND TOPICS
- 19 Image-Based Modeling and Rendering
 - 19.1 Visual Hulls
 - 19.1.1 Main Elements of the Visual Hull Model
 - 19.1.2 Tracing Intersection Curves
 - 19.1.3 Clipping Intersection Curves
 - 19.1.4 Triangulating Cone Strips
 - 19.1.5 Results
 - 19.1.6 Going Further: Carved Visual Hulls
 - 19.2 Patch-Based Multi-View Stereopsis
 - 19.2.1 Main Elements of the PMVS Model
 - 19.2.2 Initial Feature Matching
 - 19.2.3 Expansion
 - 19.2.4 Filtering
 - 19.2.5 Results
 - 19.3 The Light Field
 - 19.4 Notes
- 20 Looking at People
 - 20.1 HMM's, Dynamic Programming, and Tree-Structured Models
 - 20.1.1 Hidden Markov Models
 - 20.1.2 Inference for an HMM
 - 20.1.3 Fitting an HMM with EM
 - 20.1.4 Tree-Structured Energy Models
 - 20.2 Parsing People in Images
 - 20.2.1 Parsing with Pictorial Structure Models
 - 20.2.2 Estimating the Appearance of Clothing

<<计算机视觉>>

- 20.3 Tracking People
 - 20.3.1 Why Human Tracking Is Hard
 - 20.3.2 Kinematic Tracking by Appearance
 - 20.3.3 Kinematic Human Tracking Using Templates
- 20.4 3D from 2D: Lifting
 - 20.4.1 Reconstruction in an Orthographic View
 - 20.4.2 Exploiting Appearance for Unambiguous Reconstructions
 - 20.4.3 Exploiting Motion for Unambiguous Reconstructions
- 20.5 Activity Recognition
 - 20.5.1 Background: Human Motion Data
 - 20.5.2 Body Configuration and Activity Recognition
 - 20.5.3 Recognizing Human Activities with Appearance
 - 20.5.4 Recognizing Human Activities with Compositional Models
- 20.6 Resources
- 20.7 Notes
- 21 Image Search and Retrieval
 - 21.1 The Application Context
 - 21.1.1 Applications
 - 21.1.2 User Needs
 - 21.1.3 Types of Image Query
 - 21.1.4 What Users Do with Image Collections
 - 21.2 Basic Technologies from Information Retrieval
 - 21.2.1 Word Counts
 - 21.2.2 Smoothing Word Counts
 - 21.2.3 Approximate Nearest Neighbors and Hashing
 - 21.2.4 Ranking Documents
 - 21.3 Images as Documents
 - 21.3.1 Matching Without Quantization
 - 21.3.2 Ranking Image Search Results
 - 21.3.3 Browsing and Layout
 - 21.3.4 Laying Out Images for Browsing
 - 21.4 Predicting Annotations for Pictures
 - 21.4.1 Annotations from Nearby Words
 - 21.4.2 Annotations from the Whole Image
 - 21.4.3 Predicting Correlated Words with Classifiers
 - 21.4.4 Names and Faces
 - 21.4.5 Generating Tags with Segments
 - 21.5 The State of the Art of Word Prediction
 - 21.5.1 Resources
 - 21.5.2 Comparing Methods
 - 21.5.3 Open Problems
 - 21.6 Notes
- VII BACKGROUND MATERIAL
- 22 Optimization Techniques

<<计算机视觉>>

22.1 Linear Least-Squares Methods

22.1.1 Normal Equations and the Pseudoinverse

22.1.2 Homogeneous Systems and Eigenvalue Problems

22.1.3 Generalized Eigenvalues Problems

22.1.4 An Example: Fitting a Line to Points in a Plane

22.1.5 Singular Value Decomposition

22.2 Nonlinear Least-Squares Methods

22.2.1 Newton ' s Method: Square Systems of Nonlinear Equations.

22.2.2 Newton ' s Method for Overconstrained Systems

22.2.3 The Gauss—Newton and Levenberg—Marquardt Algorithms

22.3 Sparse Coding and Dictionary Learning

22.3.1 Sparse Coding

22.3.2 Dictionary Learning

22.3.3 Supervised Dictionary Learning

22.4 Min-Cut/Max-Flow Problems and Combinatorial Optimization

22.4.1 Min-Cut Problems

22.4.2 Quadratic Pseudo-Boolean Functions

22.4.3 Generalization to Integer Variables

22.5 Notes

Bibliography

Index

List of Algorithms

章节摘录

版权页：插图： Inference from Shading Registered images are not essential for radiometric calibration. For example, it is sufficient to have two images where we believe the histogram of E_{ij} values is the same (Grossberg and Nayar 2002). This occurs, for example, when the images are of the same scene, but are not precisely registered. Patterns of intensity around edges also can reveal calibration (Lin et al. 2004). There has not been much recent study of lightness constancy algorithms. The basic idea is due to Land and McCann (1971). Their work was formalized for the computer vision community by Horn (1974). A variation on Horn's algorithm was constructed by Blake (1985). This is the lightness algorithm we describe. It appeared originally in a slightly different form, where it was called the Retinex algorithm (Land and McCann 1971). Retinex was originally intended as a color constancy algorithm. It is surprisingly difficult to analyze (Brainard and Wandell 1986). Retinex estimates the log-illumination term by subtracting the log-albedo from the log-intensity. This has the disadvantage that we do not impose any structural constraints on illumination. This point has largely been ignored, because the main focus has been on albedo estimates. However, albedo estimates are likely to be improved by balancing violations of albedo constraints with those of illumination constraints. Lightness techniques are not as widely used as they should be, particularly given that there is some evidence that they produce useful information on real images (Brelstaff and Blake 1987). Classifying illumination versus albedo simply by looking at the magnitude of the gradient is crude, and ignores important cues. Sharp shading changes occur at shadow boundaries or normal discontinuities, but using chromaticity (Funt et al. 1992) or multiple images under different lighting conditions (Weiss 2001) yields improved estimates. One can learn to distinguish illumination from albedo (Freeman et al. 2000). Discriminative methods to classify edges into albedo or shading help (Tappen et al. 2006b) and chromaticity cues can contribute (Farenzena and Fusiello 2007).

<<计算机视觉>>

编辑推荐

《计算机视觉:一种现代方法(第2版)(英文版)》可作为高等院校计算几何、计算机图形学、图像处理、机器人学等专业学生的教材,也可供相关的专业人士阅读。

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>