IMAGE PROCESSING AND PATTERN RECOGNITION

Fundamentals and Techniques

FRANK Y. SHIH

WILEY

IEEE
IEEE PRESS

IMAGE PROCESSING AND PATTERN RECOGNITION

Fundamentals and Techniques

FRANK Y. SHIH

IEEE
IEEE PRESS

WILEY
A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data is Available

Shih, Frank Y.
Image processing and pattern recognition : fundamentals and techniques / Frank Shih.
p. cm.
ISBN 978-0-470-40461-4 (cloth)
1. Image processing. 2. Signal processing. 3. Pattern recognition systems. I. Title.
TA1637.S4744 2010
621.36′7–dc22 2009035856

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

IMAGE PROCESSING AND PATTERN RECOGNITION

IEEE Press
445 Hoes Lane
Piscataway, NJ 08854

Reviewers
Tim Newman
Ed Wong

CONTENTS

PART I
FUNDAMENTALS

1 *INTRODUCTION* 3

1.1 The World of Signals 4
1.1.1 One-Dimensional Signals 4
1.1.2 Two-Dimensional Signals 5
1.1.3 Three-Dimensional Signals 5
1.1.4 Multidimensional Signals 6
1.2 Digital Image Processing 6
1.3 Elements of an Image Processing System 11
Appendix 1.A Selected List of Books on Image Processing and Computer Vision from Year 2000 12
1.A.1 Selected List of Books on Signal Processing from Year 2000 14
1.A.2 Selected List of Books on Pattern Recognition from Year 2000 15
References 15

2 *MATHEMATICAL PRELIMINARIES* 17

2.1 Laplace Transform 17
2.1.1 Properties of Laplace Transform 19
2.2 Fourier Transform 23
2.2.1 Basic Theorems 24
2.2.2 Discrete Fourier Transform 26
2.2.3 Fast Fourier Transform 28
2.3 Z-Transform 30
2.3.1 Definition of Z-Transform 31
2.3.2 Properties of Z-Transform 32
2.4 Cosine Transform 32
2.5 Wavelet Transform 34
References 38

3 *IMAGE ENHANCEMENT* 40

3.1 Grayscale Transformation 41
3.2 Piecewise Linear Transformation 42
3.3 Bit Plane Slicing 45
3.4 Histogram Equalization 45
3.5 Histogram Specification 49
3.6 Enhancement by Arithmetic Operations 51
3.7 Smoothing Filter 52

3.8 Sharpening Filter 55
3.9 Image Blur Types and Quality Measures 59
References 61

4 *MATHEMATICAL MORPHOLOGY* 63

4.1 Binary Morphology 64
4.1.1 Binary Dilation 64
4.1.2 Binary Erosion 66
4.2 Opening and Closing 68
4.3 Hit-or-Miss Transform 69
4.4 Grayscale Morphology 71
4.4.1 Grayscale Dilation and Erosion 71
4.4.2 Grayscale Dilation Erosion Duality Theorem 75
4.5 Basic Morphological Algorithms 76
4.5.1 Boundary Extraction 76
4.5.2 Region Filling 77
4.5.3 Extraction of Connected Components 77
4.5.4 Convex Hull 78
4.5.5 Thinning 80
4.5.6 Thickening 81
4.5.7 Skeletonization 82
4.5.8 Pruning 84
4.5.9 Morphological Edge Operator 85
4.5.9.1 The Simple Morphological Edge Operators 85
4.5.9.2 Blur-Minimum Morphological Edge Operator 87
4.6 Morphological Filters 88
4.6.1 Alternating Sequential Filters 89
4.6.2 Recursive Morphological Filters 90
4.6.3 Soft Morphological Filters 94
4.6.4 Order-Statistic Soft Morphological (OSSM) Filters 99
4.6.5 Recursive Soft Morphological Filters 102
4.6.6 Recursive Order-Statistic Soft Morphological Filters 104
4.6.7 Regulated Morphological Filters 106
4.6.8 Fuzzy Morphological Filters 109
References 114

5 *IMAGE SEGMENTATION* 119

5.1 Thresholding 120
5.2 Object (Component) Labeling 122
5.3 Locating Object Contours by the Snake Model 123
5.3.1 The Traditional Snake Model 124
5.3.2 The Improved Snake Model 125
5.3.3 The Gravitation External Force Field and The Greedy Algorithm 128
5.3.4 Experimental Results 129
5.4 Edge Operators 130
5.5 Edge Linking by Adaptive Mathematical Morphology 137
5.5.1 The Adaptive Mathematical Morphology 138
5.5.2 The Adaptive Morphological Edge-Linking Algorithm 140
5.5.3 Experimental Results 141

5.6 Automatic Seeded Region Growing 146
5.6.1 Overview of the Automatic Seeded Region Growing Algorithm 146
5.6.2 The Method for Automatic Seed Selection 148
5.6.3 The Segmentation Algorithm 150
5.6.4 Experimental Results and Discussions 153
5.7 A Top-Down Region Dividing Approach 158
5.7.1 Introduction 159
5.7.2 Overview of the TDRD-Based Image Segmentation 159
5.7.2.1 Problem Motivation 159
5.7.2.2 The TDRD-Based Image Segmentation 161
5.7.3 The Region Dividing and Subregion Examination Strategies 162
5.7.3.1 Region Dividing Procedure 162
5.7.3.2 Subregion Examination Strategy 166
5.7.4 Experimental Results 167
5.7.5 Potential Applications in Medical Image Analysis 173
5.7.5.1 Breast Boundary Segmentation 173
5.7.5.2 Lung Segmentation 174
References 175

6 *DISTANCE TRANSFORMATION AND SHORTEST PATH PLANNING* 179

6.1 General Concept 180
6.2 Distance Transformation by Mathematical Morphology 184
6.3 Approximation of Euclidean Distance 186
6.4 Decomposition of Distance Structuring Element 188
6.4.1 Decomposition of City-Block and Chessboard Distance Structuring Elements 189
6.4.2 Decomposition of the Euclidean Distance Structuring Element 190
6.4.2.1 Construction Procedure 190
6.4.2.2 Computational Complexity 192
6.5 The 3D Euclidean Distance 193
6.5.1 The 3D Volumetric Data Representation 193
6.5.2 Distance Functions in the 3D Domain 193
6.5.3 The 3D Neighborhood in the EDT 194
6.6 The Acquiring Approaches 194
6.6.1 Acquiring Approaches for City-Block and Chessboard Distance Transformations 195
6.6.2 Acquiring Approach for Euclidean Distance Transformation 196
6.7 The Deriving Approaches 198
6.7.1 The Fundamental Lemmas 198
6.7.2 The Two-Scan Algorithm for EDT 200
6.7.3 The Complexity of the Two-Scan Algorithm 203
6.8 The Shortest Path Planning 203
6.8.1 A Problematic Case of Using the Acquiring Approaches 204
6.8.2 Dynamically Rotational Mathematical Morphology 205
6.8.3 The Algorithm for Shortest Path Planning 206
6.8.4 Some Examples 207
6.9 Forward and Backward Chain Codes for Motion Planning 209
6.10 A Few Examples 213
References 217

7 *IMAGE REPRESENTATION AND DESCRIPTION* 219

7.1 Run-Length Coding 219
7.2 Binary Tree and Quadtree 221
7.3 Contour Representation 223
7.3.1 Chain Code and Crack Code 224
7.3.2 Difference Chain Code 226
7.3.3 Shape Signature 227
7.3.4 The Mid-Crack Code 227
7.4 Skeletonization by Thinning 233
7.4.1 The Iterative Thinning Algorithm 234
7.4.2 The Fully Parallel Thinning Algorithm 235
7.4.2.1 Definition of Safe Point 236
7.4.2.2 Safe Point Table 239
7.4.2.3 Deletability Conditions 239
7.4.2.4 The Fully Parallel Thinning Algorithm 243
7.4.2.5 Experimental Results and Discussion 243
7.5 Medial Axis Transformation 244
7.5.1 Thick Skeleton Generation 252
7.5.1.1 The Skeleton from Distance Function 253
7.5.1.2 Detection of Ridge Points 253
7.5.1.3 Trivial Uphill Generation 253
7.5.2 Basic Definitions 254
7.5.2.1 Base Point 254
7.5.2.2 Apex Point 254
7.5.2.3 Directional Uphill Generation 254
7.5.2.4 Directional Downhill Generation 255
7.5.3 The Skeletonization Algorithm and Connectivity Properties 256
7.5.4 A Modified Algorithm 259
7.6 Object Representation and Tolerance 260
7.6.1 Representation Framework: Formal Languages and Mathematical Morphology 261
7.6.2 Dimensional Attributes 262
7.6.2.1 The 2D Attributes 262
7.6.2.2 The 3D Attributes 263
7.6.2.3 Tolerancing Expression 263
References 265

8 *FEATURE EXTRACTION* 269

8.1 Fourier Descriptor and Moment Invariants 269
8.2 Shape Number and Hierarchical Features 274
8.2.1 Shape Number 274
8.2.2 Significant Points Radius and Coordinates 276
8.2.3 Localization by Hierarchical Morphological Band-Pass Filter 277
8.3 Corner Detection 278
8.3.1 Asymmetrical Closing for Corner Detection 280
8.3.2 Regulated Morphology for Corner Detection 281
8.3.3 Experimental Results 283
8.4 Hough Transform 286
8.5 Principal Component Analysis 289

8.6 Linear Discriminate Analysis 291
8.7 Feature Reduction in Input and Feature Spaces 293
8.7.1 Feature Reduction in the Input Space 293
8.7.2 Feature Reduction in the Feature Space 297
8.7.3 Combination of Input and Feature Spaces 299
References 302

9 *PATTERN RECOGNITION* 306
9.1 The Unsupervised Clustering Algorithm 307
9.1.1 Pass 1: Cluster's Mean Vector Establishment 308
9.1.2 Pass 2: Pixel Classification 309
9.2 Bayes Classifier 310
9.3 Support Vector Machine 313
9.3.1 Linear Maximal Margin Classifier 313
9.3.2 Linear Soft Margin Classifier 315
9.3.3 Nonlinear Classifier 316
9.3.4 SVM Networks 317
9.4 Neural Networks 320
9.4.1 Programmable Logic Neural Networks 321
9.4.2 Pyramid Neural Network Structure 323
9.4.3 Binary Morphological Operations by Logic Modules 324
9.4.4 Multilayer Perceptron as Processing Modules 327
9.5 The Adaptive Resonance Theory Network 334
9.5.1 The ART1 Model and Learning Process 334
9.5.2 The ART2 Model 337
9.5.2.1 Learning in the ART2 Model 337
9.5.2.2 Functional-Link Net Preprocessor 339
9.5.3 Improvement of ART Model 341
9.5.3.1 Problem Analysis 341
9.5.3.2 An Improved ART Model for Pattern Classification 342
9.5.3.3 Experimental Results of the Improved Model 344
9.6 Fuzzy Sets in Image Analysis 346
9.6.1 Role of Fuzzy Geometry in Image Analysis 346
9.6.2 Definitions of Fuzzy Sets 347
9.6.3 Set Theoretic Operations 348
References 349

PART II
APPLICATIONS

10 *FACE IMAGE PROCESSING AND ANALYSIS* 355
10.1 Face and Facial Feature Extraction 356
10.1.1 Face Extraction 357
10.1.2 Facial Feature Extraction 362
10.1.3 Experimental Results 367
10.2 Extraction of Head and Face Boundaries and Facial Features 370
10.2.1 The Methodology 372

10.2.1.1 Smoothing and Thresholding 372
10.2.1.2 Tracing Head and Face Boundaries 374
10.2.1.3 Locate Facial Features 374
10.2.1.4 Face Boundary Repairing 374
10.2.2 Finding Facial Features Based on Geometric Face Model 375
10.2.2.1 Geometric Face Model 375
10.2.2.2 Geometrical Face Model Based on Gabor Filter 377
10.2.3 Experimental Results 378
10.3 Recognizing Facial Action Units 378
10.3.1 Facial Action Coding System and Expression Database 379
10.3.2 The Proposed System 382
10.3.3 Experimental Results 383
10.4 Facial Expression Recognition in JAFFE Database 386
10.4.1 The JAFFE Database 388
10.4.2 The Proposed Method 389
10.4.2.1 Preprocessing 389
10.4.2.2 Feature Extraction 389
10.4.2.3 Expression Classification 390
10.4.3 Experimental Results and Performance Comparisons 390
References 392

11 *DOCUMENT IMAGE PROCESSING AND CLASSIFICATION* 397

11.1 Block Segmentation and Classification 398
11.1.1 An Improved Two-Step Algorithm for Block Segmentation 399
11.1.2 Rule-Based Block Classification 400
11.1.3 Parameters Adaptation 402
11.1.4 Experimental Results 403
11.2 Rule-Based Character Recognition System 407
11.3 Logo Identification 411
11.4 Fuzzy Typographical Analysis for Character Preclassification 414
11.4.1 Character Typographical Structure Analysis 415
11.4.2 Baseline Detection 416
11.4.3 Tolerance Analysis 417
11.4.4 Fuzzy Typographical Categorization 419
11.4.5 Experimental Results 424
11.5 Fuzzy Model for Character Classification 426
11.5.1 Similarity Measurement 427
11.5.2 Statistical Fuzzy Model for Classification 430
11.5.3 Similarity Measure in Fuzzy Model 433
11.5.4 Matching Algorithm 434
11.5.5 Classification Hierarchy 436
11.5.6 Preclassifier for Grouping the Fuzzy Prototypes 437
11.5.7 Experimental Results 439
References 441

12 *IMAGE WATERMARKING* 444

12.1 Watermarking Classification 445
12.1.1 Blind Versus Non Blind 445
12.1.2 Perceptible Versus Imperceptible 446

12.1.3 Private Versus Public 446
12.1.4 Robust Versus Fragile 446
12.1.5 Spatial Domain Versus Frequency Domain 447
12.2 Spatial Domain Watermarking 448
12.2.1 Substitution Watermarking in the Spatial Domain 448
12.2.2 Additive Watermarking in the Spatial Domain 450
12.3 Frequency-Domain Watermarking 452
12.3.1 Substitution Watermarking in the Frequency Domain 452
12.3.2 Multiplicative Watermarking in the Frequency Domain 453
12.3.3 Watermarking Based on Vector Quantization 455
12.3.4 Rounding Error Problem 456
12.4 Fragile Watermark 458
12.4.1 The Block-Based Fragile Watermark 458
12.4.2 Weakness of the Block-Based Fragile Watermark 459
12.4.3 The Hierarchical Block-Based Fragile Watermark 460
12.5 Robust Watermark 461
12.5.1 The Redundant Embedding Approach 461
12.5.2 The Spread Spectrum Approach 462
12.6 Combinational Domain Digital Watermarking 462
12.6.1 Overview of Combinational Watermarking 463
12.6.2 Watermarking in the Spatial Domain 464
12.6.3 The Watermarking in the Frequency Domain 465
12.6.4 Experimental Results 466
12.6.5 Further Encryption of Combinational Watermarking 470
References 471

13 *IMAGE STEGANOGRAPHY* 474

13.1 Types of Steganography 476
13.1.1 Technical Steganography 476
13.1.2 Linguistic Steganography 477
13.1.3 Digital Steganography 478
13.2 Applications of Steganography 478
13.2.1 Covert Communication 478
13.2.2 One-Time Pad Communication 479
13.3 Embedding Security and Imperceptibility 480
13.4 Examples of Steganography Software 480
13.4.1 S-Tools 481
13.4.2 StegoDos 481
13.4.3 EzStego 481
13.4.4 JSteg-Jpeg 481
13.5 Genetic Algorithm-Based Steganography 482
13.5.1 Overview of the GA-Based Breaking Methodology 482
13.5.2 The GA-Based Breaking Algorithm on SDSS 485
13.5.2.1 Generating the Stego Image on the Visual Steganalytic System 486
13.5.2.2 Generating the Stego Image on the IQM-Based Steganalytic System (IQM-SDSS) 486
13.5.3 The GA-Based Breaking Algorithm on FDSS 487
13.5.4 Experimental Results 489

13.5.4.1 The GA-Based Breaking Algorithm on VSS 489
13.5.4.2 The GA-Based Breaking Algorithm on IQM-SDSS 490
13.5.4.3 The GA-Based Breaking Algorithm on JFDSS 491
13.5.5 Complexity Analysis 493
References 494

14 *SOLAR IMAGE PROCESSING AND ANALYSIS* 496

14.1 Automatic Extraction of Filaments 496
14.1.1 Local Thresholding Based on Median Values 497
14.1.2 Global Thresholding with Brightness and Area Normalization 501
14.1.3 Feature Extraction 506
14.1.4 Experimental Results 511
14.2 Solar Flare Detection 515
14.2.1 Feature Analysis and Preprocessing 518
14.2.2 Classification Rates 519
14.3 Solar Corona Mass Ejection Detection 521
14.3.1 Preprocessing 523
14.3.2 Automatic Detection of CMEs 525
14.3.2.1 Segmentation of CMEs 525
14.3.2.2 Features of CMEs 525
14.3.3 Classification of Strong, Medium, and Weak CMEs 526
14.3.4 Comparisons for CME Detections 529
References 531

INDEX 535

PART I

FUNDAMENTALS

CHAPTER 1

INTRODUCTION

An image is a subset of a signal. A *signal* is a function that conveys information generally about the behavior of a physical system or attributes of some phenomenon. A simple example is the traffic signal that uses three universal color codes (red, yellow, and green) signaling the moment to stop, drive, or walk. Although signals can be represented in many ways, in all cases the information is contained in a pattern of variations of some form, and with that information is typically transmitted and received over a medium. Electrical quantities such as current and voltage are called *electrical signals*, which are often used in radio, radar, sonar, telephone, television, and many other areas. An acoustic wave signal can convey speech or music information, in which people often speak of a strong or weak signal when the sound is referred to its clarity and audibility. A thermocouple can convey temperature, and a pH meter can convey the acidity of a solution.

A signal may take a form of time variations or a spatially varying pattern. Mathematically speaking, signals are represented as functions of one or more independent variables that can be either continuous or discrete. *Continuous-time* signals are defined at a continuum of the time variable. *Discrete-time* signals are defined at discrete instants of time. *Digital* signals are those for which both time and amplitude are discrete. The continuous-time and continuous-amplitude signals are called *analog* signals. Analog signals that have been converted to digital forms can be processed by a computer or other digital devices.

Signal processing is the process of extracting information from the signal. *Digital signal processing* (DSP) is concerned with the representation of signals by sequences of numbers or symbols and processing of these sequences. It was initiated in the seventeenth century and has become an important modern tool in the tremendously diverse fields of science and technology. The purpose of such processing is to estimate characteristic parameters of a signal or to transform a signal into a form that is more sensible to human beings. DSP includes subfields such as digital image processing, video processing, statistical signal processing, signal processing for communications, biomedical signal processing, audio and speech signal processing, sonar and radar signal processing, sensor array processing, spectral estimation, and so on.

Human beings possess a natural signal processing system. "Seeing" takes place in the visual system and "hearing" takes place in the auditory system. Human visual system (HVS) plays an important role in navigation, identification, verification, gait, gesture, posture, communication, psychological interpretation, and so on. Human

Image Processing and Pattern Recognition by Frank Shih

auditory system converts sound waves into nerve impulses, to analyze auditory events, remember and recognize sound sources, and perceive acoustic sequences. As the speed, capability, and economic advantages of modern signal processing devices continue to increase, there is simultaneously an increase in efforts aimed at developing sophisticated, real-time automatic systems capable of emulating human abilities. Because of digital revolution, digital signals have been increasingly used. Most household electronic devices are based entirely or almost entirely upon digital signals. The entire Internet is a network of digital signals, as is modern mobile phone communication.

1.1 THE WORLD OF SIGNALS

The world is filled with many kinds of signals; each has its own physical meaning. Sometimes the human body is incapable of receiving a special signal or interpreting (decoding) a signal, so the information that the signal intends to convey cannot be captured. Those signals are not to be said nonsense or insignificant, but conversely they are exactly what people are working very hard to understand. The more we learn from the world's signals, the better living environment we can provide. Furthermore, some disaster or damage can be avoided if a warning signal can be sensed in advance. For example, it was recorded historically that animals, including rats, snakes, and weasels, deserted the Greek city of Helice in droves just days before a quake devastated the city in 373 B.C. Numerous claims have been made that dogs and cats usually behave strangely before earthquake by barking, whining, or showing signs of nervousness and restlessness.

The characteristics of a signal may be one of a broad range of shapes, amplitudes, time durations, and perhaps other physical properties. Based on the sampling of time axis, signals can be divided into continuous-time and discrete-time signals. Based on the sampling of time and amplitude axes, signals can be divided into analog and digital signals. If signals repeat in some period, they are called periodic signals; otherwise, aperiodic or nonperiodic signals. If each value of a signal is fixed by a mathematical function, it is called a deterministic signal; otherwise, a random signal that has uncertainty about its behavior. In the category of dimensionality, signals are divided into one-dimensional (1D), two-dimensional (2D), three-dimensional (3D), and multidimensional signals, which are further explained below.

1.1.1 One-Dimensional Signals

A 1D signal is usually modeled as an ensemble of time waveforms, for example, $x(t)$ or $f(t)$. One-dimensional signal processing has a rich history, and its importance is evident in such diverse fields as biomedical engineering, acoustics (Beranek, 2007), sonar (Sun et al., 2004), radar (Gini et al., 2001), seismology (Al-Alaoui, 2001), speech communication, and many others. When we use a telephone, our voice is converted to an electrical signal and through telecommunication systems circulates around the Earth. The radio signals, which are propagated through free space and by radio receivers, are converted into sound. In speech transmission and recognition, one