MATRIX STRUCTURED IMAGE PROCESSING

EDWARD R. DOUGHERTY
CHARLES R. GIARDINA

MATRIX STRUCTURED IMAGE PROCESSING

EDWARD R. DOUGHERTY, Ph.D.
Fairleigh Dickinson University

CHARLES R. GIARDINA, Ph.D.
Stevens Institute of Technology

Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632

Library of Congress Cataloging-in-Publication Data

Dougherty, Edward R. (date)
Matrix structured image processing.

Includes index.
1. Image processing—Digital techniques.
2. Computer vision. I. Giardina, Charles Robert.
II. Title.
TA1632.G53 1987 006.3'7 86-12374
ISBN 0-13-565623-0

Editorial/production supervision and
interior design: Gloria Jordan
Cover design: George Cornell
Manufacturing buyer: Ed O'Dougherty

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-565623-0 025

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Prentice-Hall of Southeast Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

PREFACE

Because of the necessity to provide computers with the ability to make "intelligent" decisions concerning visual data, the relatively new field of digital image processing has undergone a substantial growth spurt during the past twenty years. As is often the case in the scientific quest for new knowledge, the organization of that knowledge and its concomitant accessibility to investigators in related fields receives insufficient attention. It is our hope that this book will help rectify the situation.

It is easy to put into succinct terms the rationale behind our approach: to embed existing digital imaging techniques into a consistent and unified mathematical framework, and to accomplish this task at a level accessible to upper division undergraduates and beginning graduate students in computer science, engineering, the physical sciences, and those other areas of contemporary technology where computer-assisted analysis of digital images plays a central role.

To reach this goal, we introduce at the outset a mathematical structure, to be employed throughout the text, for the representation of digital images. This structure, the "bound matrix," is used in conjunction with block diagrams to serve as a vehicle for the concise expression of digital image processing operators. While we do not present rigorous mathematical proofs, the definitions are precise and the properties given are exactly expressed so that implementation and application can be accomplished in a systematic fashion.

Throughout the text there is an abundance of illustrative examples to help students "walk through" the imaging algorithms. We believe that methodology must always remain in a "how to" mode, allowing students to gain a working knowledge of fundamental image processing procedures. Our desire was not to write an encyclopedia of imaging operations; rather, we wanted to give a representative selection of techniques at an implementable level.

The concept of a digital image, together with its bound matrix representation, is given in Chapter 1. That chapter also includes primitive mathematical operations on bound matrices from which digital image processing algorithms are composed. Fundamental gray-level processing procedures are discussed in Chapter 2.

Edge detection is introduced in Chapter 3. The emphasis is on the Prewitt, Sobel, Roberts and other gradient-type detectors. Compass gradient techniques such as the Kirsh are also presented.

In Chapter 4, the elements of morphology are presented. Because of its intuitive geometric character and the relatively fast processing speed of its underlying operators, morphology is becoming evermore important in digital image processing. This chapter also details basic Minkowski algebra as well as its essential properties. It also introduces size distributions, morphological covariance, and the skeleton.

Versions of the well-known tracking, region growing, curve filling, and template matching methodologies are given in Chapter 5. These are presented in terms of precise specifications within the bound matrix structure. Detailed walk-throughs are provided.

A major topic in any discussion of image processing is transforms. The concept of a matrix image transform is introduced in Chapter 6 for a systematic presentation. Among the topics covered are the DFT, cyclic convolutions, the Hadamard transform, and the singular value decomposition transform.

The last chapter presents novel computer architectures well-suited to the natural structure of digital images and to operations on those images. Data flow, systolic and wavefront array systems are explained. Matrix multiplication, which plays a crucial role in image transform methodology, is implemented using both a systolic system and a wavefront array.

Insofar as prerequisites for the text are concerned, they are minimal. Except for some knowledge of matrix theory, no specific mathematical, engineering, or computer science background is required. Nonetheless, some sophistication with regard to algorithm development and computer implementation would be helpful.

Edward R. Dougherty
Fairleigh Dickinson University

Charles R. Giardina
Stevens Institute of Technology

CONTENTS

PREFACE ix

INTRODUCTION 1

What Does It Mean to Form an Image? 1
Four Levels of Image Processing 3
Computer Vision 4

1 IMAGE-PROCESSING OPERATIONS 9

1.1 Digital Images 9
1.2 Representation of Digital Images 10
1.3 Bound Matrices 13
1.4 Information Density of a Bound Matrix 14
1.5 Block Diagrams 15
1.6 Arithmetic-type Binary Operations for Images 17
1.7 Unary Arithmetic Operators for Images 20
1.8 Translation Operation 21
1.9 Rotation and Flip Operations 22
1.10 Data-base-type Image Operations 25
1.11 Structural Transformations 28
1.12 Matrix Multiplication 33
Exercises 34

2 GRAY-LEVEL PROCESSING 37

2.1 Thresholding 37
2.2 Smoothing 44
2.3 Moving-average Filters 53
2.4 A Local Noise-removal Technique 55
2.5 Gray-level Histogram 59
2.6 Histogram Equalization 63
2.7 Co-occurrence Matrix 65

3 EDGE DETECTION 72

3.1 Concept of an Edge 72
3.2 Partial Difference Operators 74
3.3 Gradient 77
3.4 Gradient-type Edge Detectors 83
3.5 Compass Gradients 87
3.6 Textural Edge Detection 90

4 MORPHOLOGY 98

4.1 Definition 98
4.2 Constant Images 98
4.3 Minkowski Algebra 101
4.4 Opening and Closing 105
4.5 Properties 108
4.6 Granulometries 112
4.7 Size Distribution by Linear Erosion 117
4.8 Morphological Covariance 118
4.9 Skeleton 121

5 TOPOLOGICAL OPERATIONS 131

5.1 Vertical Tracking 131
5.2 Region Growing 140
5.3 Curve Filling 149
5.4 Template Matching 154

6 TRANSFORM TECHNIQUES 168

6.1 Image-processing Transform Techniques 168
6.2 Transform Techniques Widely Prevalent in Image Processing 170
6.3 Vector Transform 171
6.4 Least-squares Image Restoration 172
6.5 Operations on Bound Matrices Similar to Regular Matrices 179
6.6 Matrix Image Transforms 185
6.7 Notation When Using the Matrix Image Transform 189
6.8 Discrete Fourier Transform 192
6.9 Inverse Discrete Fourier Transform 194
6.10 Cyclic Convolutions 195
6.11 Vector Transformation Method for Finding Circular Convolutions 198
6.12 Discrete Fourier Transform for Determining Cyclic Convolutions 202
6.13 Cyclic Correlation and the Discrete Fourier Transform 204
6.14 Discrete Cosine Transform 205
6.15 Hadamard Transform 207
6.16 Singular-value Decomposition Transform 209
6.17 Singular-value Decomposition for Least-squares Problems 214
6.18 Outer Product Representation of the Singular-value Decomposition 218

7 IMAGING ARCHITECTURE 223

7.1 Data-flow Systems 223
7.2 Systolic Processors 226
7.3 Systolic Matrix Multiplication Using Block Diagrams 232
7.4 Wavefront Array Processors 235
7.5 Describing Wavefront Array Matrix Multiplication Using Block Diagrams 237

TABLE OF OPERATORS 243

APPENDIX 245

INDEX 253

INTRODUCTION

What Does It Mean to Form an Image?

Perhaps the best way to understand the problems and goals that constitute the subject of image processing is to reflect on our own human vision system. Although visual sensation, object recognition, and the decision for action seem almost instantaneous, the complexity of the process is to this day beyond full comprehension. Imagine a batter awaiting the pitch that will momentarily arrive from the mound. In a fraction of a second, a stringent real-time constraint, the following must be accomplished:

1) Visual sensors, the batter's eyes, must *sense* the reflected light coming from the ball.
2) This purely physical sensory data must be organized into an *image*; the image must be *created.*
3) The image must be *cleaned up*; it must be preprocessed so as to yield an image that is not distorted by the act of sensation. The image must be *restored* to some "perfect" image that, by experience, the batter believes it to be. For example, he expects a ball to be coming toward him, and he must restore the image so that *noisy* interference such as light or fog do not result in something other than a ball.
4) The image is *enhanced* so that the ball stands out more starkly from the background.

5) The enhanced image is *segmented*; the ball is cut from the rest of the image except insofar as the background is necessary to determine the speed and trajectory of the ball.
6) Fundamental *features* of the image are *selected* and perhaps computed. *Compression* takes place in that extraneous information, that which is not required for the task at hand (hitting the ball), is filtered out and removed from further processing by the brain.
7) The position and internal data of the batter himself are *registered* and *linked* to the image. The complexity of this step is extreme in that the new knowledge concerning the image must be combined with the batter's self-knowledge of himself, including his own inertial system.
8) The total image must be *classified*; for example, the type of pitch, such as a curve ball, fast ball, or slider, must be determined.
9) A *decision*, or a complex of decisions, must be made, for example, to swing or not to swing, where to swing, and how hard to swing.

To fully appreciate the enormity of the task faced by any vision system, one should recognize that each of these stages consists of numerous substages, all to be accomplished in less than a second!

The purpose and overall logical structure of a computer vision system are essentially the same as for the human vision system; indeed, our approach to the problem of artificial vision systems naturally follows the conception we have of our own vision system. It could not be otherwise, for the manner in which we frame our understanding of optical intelligence can only arise from reflection on and investigation of that unique system in which vision and human intelligence are linked, the human vision system. Image understanding, as it applies to artificial systems, must utilize the notions of shape, texture, edge, and position, among others, that are common to human comprehension. It is precisely these categories to which we refer when we make decisions based on optical data. Not only must incoming data be represented digitally in a computer vision system, but they must also be organized into configurations and categories that model those arising naturally in the human being.

At one time it was thought that the human vision system was essentially passive. Sensation was not clearly distinguished from perception. The role of the brain in the constitution of the image was overlooked, and the perceived image was thought to be a faithful copy of the "actual" image. Today we realize that the "actual" image is not a subject of scientific investigation. All systems create images from the data of sensation. While it is true that we naturally hold out our Euclidean imaging system as the standard, it is important to keep in mind that ours is only one model among many. All scientific endeavor is anthropocentric in that it takes place relative to the sense data received by human beings; nevertheless, an artificial system is artificial only insofar as it is not human. It too must be equipped with sensors, it too must process the initial image, and it too must classify and decide based on heuristic criteria established *independently* of any particular bundle of sensations.

Four Levels of Image Processing

To begin, we would like to identify four macrolevels of image processing:

0) Image representation
1) Image-to-image transformations
2) Image-to-parameter (image-to-real number) transformations
3) Parameter-to-decision transformations

At level 0, we encounter the problem of digital representation. The sensed data must be organized into a data structure compatible with the digital nature of the computer. Even if we are to work only with black and white pictures, an assumption that will be made throughout this text, human perception seemingly involves a continuously changing level of gray, whereas the computer can hold only discrete gray levels. Moreover, if we look at a two-dimensional photograph, it appears as though the space upon which the gray-level values are defined is a Euclidean plane. In other words, it appears that a two-dimensional black and white image should be represented by a function $I = f(x, y)$, where the position in the plane is given by the coordinate pair (x, y) and the gray-level intensity I is a real number.

At once two difficulties arise. First, there is the problem of *quantization*; what discrete values of I should be allowed? Second, there is the problem of *digitization*; how should we choose some finite number of coordinate pairs (x, y) at which to specify the intensities of the image? In simple terms, how should the image be created from the data so that it still maintains the optical properties necessary for an acceptable modeling of the human image? Later we shall give an account of digital image representation; however, the *sampling problem*, that is, how to choose quantization and digitization levels, is essentially beyond the level of this book.

Level 1 involves image-to-image operations, for instance, those that occur in restoration, enhancement, and segmentation. At level 1, most processing takes place, and it is with this level that most of this text is concerned. Fundamental to the implementation of the algorithms involved in image-to-image transformations is a precise, rigorously defined mathematical structure. This structure must allow the specification of unambiguous image operations, and it should be designed to facilitate the development of both imaging languages and imaging architectures. For this purpose the notion of *bound matrices* is introduced in Section 1.3, and the corresponding definitions of low-level imaging operations are developed throughout the remainder of Chapter 1.

Once an image has been prepared for the extraction of information, processing must occur that results in the establishment of parameters that give quantitative specification of the information. There may be several different parameters, or sets of parameters, each related to a particular feature of the image. We refer to an array of such parameters as a *feature vector*. In a sense, our scientific understanding of the image is encoded in the feature parameters we have extracted from the image. Operations that result in feature parameters are the level 2 transformations. Examples of such transformations are the textural descriptors of mathematical morphology, which are described in Chapter 4.

Once quantitative knowledge has been extracted from the image in the form of parameters, logical decisions can be implemented. The chain is now complete. The image, which was originally constituted from incoming sensory data, is now classified, and a decision is made as a result of this classification. Such logical transformations belong to level 3. What is desired is identification of the image as a member of a predetermined finite class of candidates; for example, the image is a square and not a triangle or circle. Based on this classification, a knowledge-based decision to choose one of a number of actions is taken. The transformations at this level are problematic in that they require a goodly amount of preprogrammed heuristics. They always involve some degree of uncertainty owing to the earlier quantization and digitization modeling and to the multiple levels of processing that have already occurred. However, keep in mind that this uncertainty is no different from the uncertainty faced by any human attempting to decide between several alternatives. That decision, too, is affected by the accuracy of the visual model, the completeness of the knowledge base, and the quality of the optical processing.

Computer Vision

Our thrust has been in the direction of recognition and decision. The goal of the text is to introduce fundamental digital techniques that lead to the formation of quantitative knowledge parameters and to quantitative decision techniques, the latter being central to the development of artificially intelligent systems. The nine stages of processing introduced metaphorically for the baseball batter constitute the stages of a *recognizer* system. An image of a three-dimensional set of objects is obtained, and the recognizer must decide whether or not some prespecified object is a member of the set. Moreover, all the processing must be accomplished in real time; that is, the absolute difference in time from when the object is sensed to when a decision is made must be less than some prespecified amount. This prespecified amount of time may change from application to application, but whatever it might be, the predetermined time line cannot be exceeded.

Whether or not a given object can be detected in real time by only processing imaging information is problematic; detection capability is a function of numerous factors. Among them are the overall set of objects that might be viewed, their similarity to the object one wishes to detect, the different possible aspect angles used in viewing the object (the more angles, the greater the difficulty of identification), the type of sensors involved, the amount of distortion and noise, the algorithms and procedures employed, the architecture of the computer being utilized, the specific real-time constraint, and the confidence level required for making a final decision.

An ideal recognizer system is depicted in Figure I.1. In an ideal system, we suppose that a set of objects is observed by a perfect imaging sensor. This sensor might utilize an optical technique such as a camera, or it might sense the heat as in an infrared (IR) sensor, or it might sense the illumination as in a synthetic aperture radar (SAR). In any event, the sensor is assumed to be perfect in that it results in an image without noise or distortion. In fact, such assumptions are not realistic. No sensor is perfect.

A more realistic scenario is depicted in Figure I.2, where one can recognize the nine stages given earlier for the batter. Usually the image is *created* by a mathematical

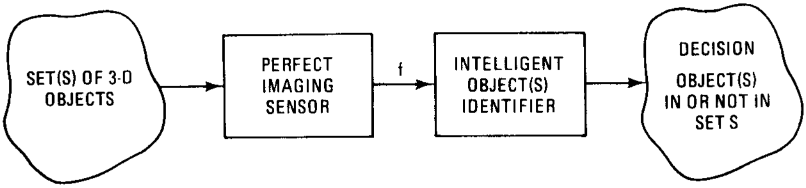

Figure I.1 Object Recognizer Using Imaging

procedure applied to the incoming sense data. It may be cleaned up, or *restored*, by processing it so as to produce an estimate of the perfect image. It may be *enhanced* to make more definitive those features of the image that are key for future processing. Once enhanced, the image is often *segmented* into mutually exclusive pieces that are easier to process and recognize in subsequent object identification procedures. Features are *selected* that are useful for classification. Since the sensor itself does not represent an absolute frame of reference, its relevant aspects, such as velocity and attitude, must be *registered*. These sensory data must be coordinated with the image data. Finally, before decision procedures can be implemented, *classification* operations must take place. For instance, the segmented objects in the image may be checked against a predetermined class of objects to see if a desired object is present, or an estimate of the location and velocity of an object within the image may be computed. In any case, a large number of steps is involved from integration of the object by the sensors to making the actual *detection decision*.

A digital computer that performs the preceding steps in real time must usually have an architecture that is custom-made or developed in accordance with the algorithms that perform the functions associated with each of these blocks. An in-depth account of relevant architectures is given in Chapter 7. For the interim, however, consider Figure I.3. In this diagram the digital processor is a multiple-instruction, multiple-data (MIMD) or a combined single-instruction, multiple-data (SIMD) and multiple-instruction, single-data (MISD) type of processor. These processors may make use of pipelining and array processing. In addition, they are configured to increase the throughput and thereby possibly meet the real-time constraints of the image-processing system.

In the succeeding chapters, we shall develop a precise structure for image representation and utilize certain fundamental imaging operations to implement higher-level procedures that accomplish many of the tasks outlined. It is not our purpose to give an encyclopedia of imaging operations, but rather to illustrate a representative collection of transformations that lead to artificially intelligent image analysis. As noted earlier, the roles of humanly understood concepts, such as shape and texture, are central, as are the procedures that result in the implementation of the stages in Figure I.2.

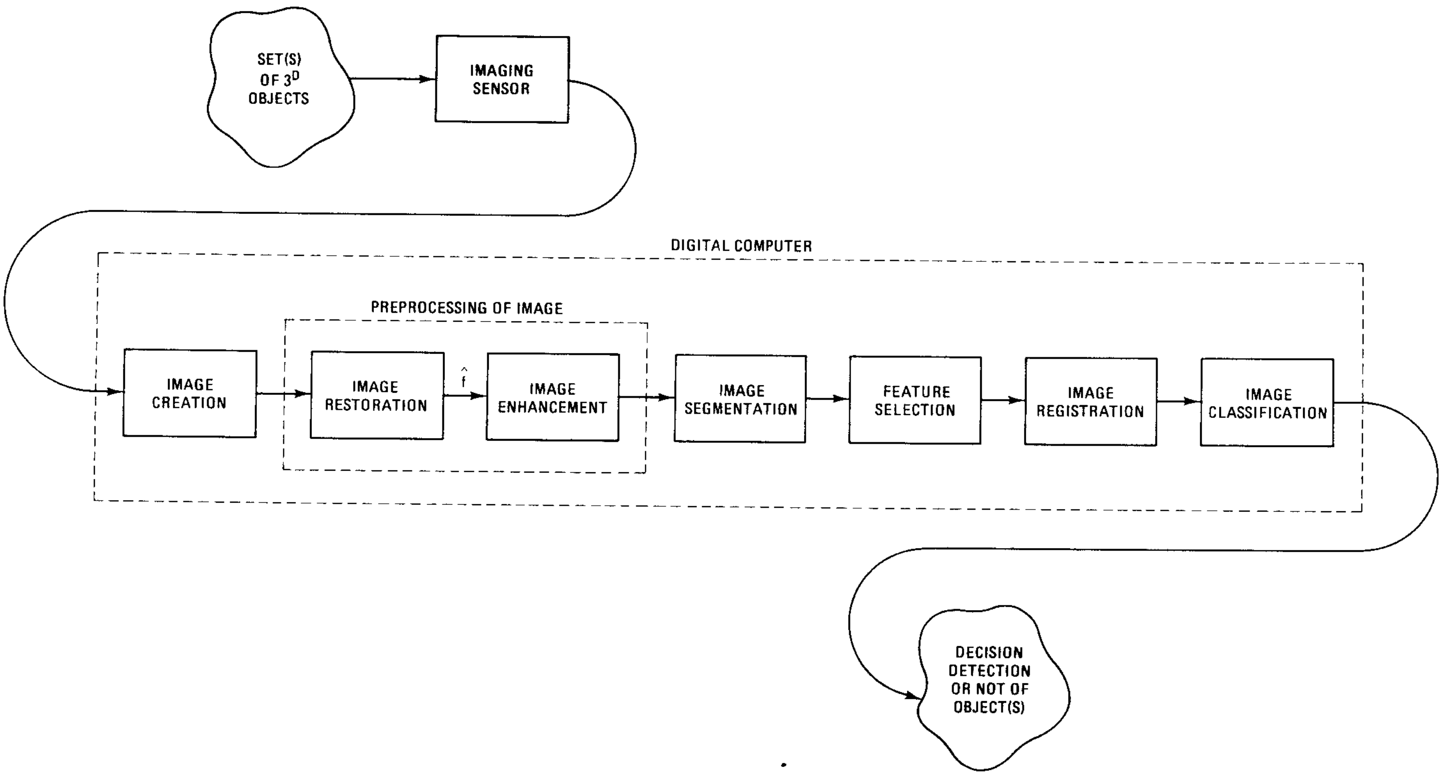

Figure I.2 Imaging Operations for Recognizer

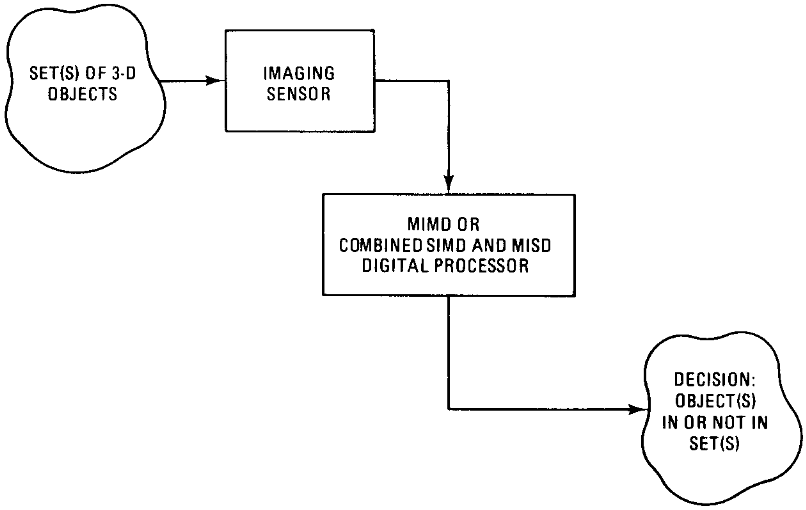

Figure I.3 Digital Object Recognizer

1

IMAGE-PROCESSING OPERATIONS

In this chapter the notion of a digital image is introduced. It is presented together with a representation structure called a *bound matrix*. The operations which are fundamental to the processing of images are defined and illustrated. These operations are the building blocks from which important image-processing transformations are formed.

1.1 Digital Images

A digital image is similar to a matrix or array of numbers. It is a useful vehicle for representing pictures and is well suited for digital machines. Before the actual definition is given, an example will be presented to furnish motivation.[1]

Consider a television set with a fixed, nonchanging square picture. A digital image representing this picture is desired. A sensor, in the form of a light meter, is provided to help in the determination of this image. The meter has integral readings from 0 to $2^n - 1$, where 0 denotes white and $2^n - 1$ denotes black. The intermediate numbers constitute a scale running from light to dark, with the light values near 0 and the dark values near $2^n - 1$. To obtain the digital image, partition the television screen into 2^n smaller squares. This might be accomplished by drawing equidistant vertical and horizontal lines on the screen. Next, on a large piece of opaque material such as cardboard, cut a square hole equal in size to a single small square on the screen. Then place the opaque material and the light meter at a fixed distance from the screen. Finally, take 2^n

[1]From G. T. Herman, *Image Reconstruction from Projection*, Academic Press, Inc., New York, 1980.