MULTIDIMENSIONAL DIGITAL SIGNAL PROCESSING

MULTIDIMENSIONAL DIGITAL SIGNAL PROCESSING

DAN E. DUDGEON
Lincoln Laboratory, M.I.T.

RUSSELL M. MERSEREAU
School of Electrical Engineering
Georgia Institute of Technology

PRENTICE-HALL, INC., Englewood Cliffs, New Jersey 07632

Library of Congress Cataloging in Publication Data

Dudgeon, Dan E.
Multidimensional digital signal processing.

(Prentice-Hall signal processing series)
Includes bibliographical references and index.
1. Signal processing—Digital techniques.
I. Mersereau, Russell M. II. Title. III. Series.
TK5102.5.D83 1984 621.38'042 83-3135
ISBN 0-13-604959-1

Editorial/production supervision and
interior design: *Ellen Denning*
Manufacturing buyer: *Anthony Caruso*

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-604959-1

PRENTICE-HALL INTERNATIONAL, INC., *London*
PRENTICE-HALL OF AUSTRALIA PTY. LIMITED, *Sydney*
EDITORA PRENTICE-HALL DO BRASIL, Ltda., *Rio de Janeiro*
PRENTICE-HALL CANADA INC., *Toronto*
PRENTICE-HALL OF INDIA PRIVATE LIMITED, *New Delhi*
PRENTICE-HALL OF JAPAN, INC., *Tokyo*
PRENTICE-HALL OF SOUTHEAST ASIA PTE. LTD., *Singapore*
WHITEHALL BOOKS LIMITED, *Wellington, New Zealand*

PREFACE

This book is the result of a suggestion by Professor Alan V. Oppenheim, who was our doctoral thesis supervisor at M.I.T. and who serves as the series editor for the Prentice-Hall series of books on signal processing, that we should write a senior- or first-year-graduate-level textbook on multidimensional digital signal processing. It is intended to be used in a one-semester course which would follow a basic course in digital signal processing using a text such as *Digital Signal Processing* by Oppenheim and Schafer (Prentice-Hall, 1975).

This text provides the student with a basic background of multidimensional signal processing theory with an emphasis on the differences and similarities between the one-dimensional and multidimensional cases. We have endeavored to write a text that will develop the student's intuition and motivation for this field without boring him or her with lengthy formal derivations, theorems, and proofs. Mathematical formality has its place, of course, but we feel that it should spring from an intuitive understanding of how things work, not the other way around. We hope the more mathematically inclined readers will understand, and benefit from, our informal approach.

There are several good books on the topic of digital image processing already available, so we have not attempted to duplicate the bulk of that subject matter in our book. Instead, we have tried to develop the theory of multidimensional signal processing, which not only serves as the foundation for image processing but also has applicability to other areas, such as array processing (e.g., radar, sonar, seismic signal processing, and radio astronomy).

In this book we assume that the reader has knowledge of one-dimensional digital signal processing theory, including linear shift-invariant systems, the discrete Fourier transform (DFT), the fast Fourier transform (FFT), linear filtering, the z-transform, stability, and power spectrum estimation. These concepts are not reviewed explicitly, but they are introduced in the two-dimensional context in a straightforward and rudimentary way.

Chapter 1 introduces the basic concepts of multidimensional signals and systems, focusing in particular on two-dimensional signals and linear shift-invariant (LSI) systems. The notion of the impulse response is introduced as one way to characterize LSI systems. The multidimensional Fourier transform is defined and used to compute the frequency response of two-dimensional LSI systems. Strategies for sampling two-dimensional continuous signals are also discussed.

In Chapter 2, the multidimensional discrete Fourier transform is introduced and algorithms for its efficient computation are presented in detail. The fast Fourier transform algorithm is shown to be applicable to signals sampled with arbitrary periodic sampling geometries. The close relationship between one-dimensional and multidimensional DFTs is also discussed.

Chapter 3 focuses on the design and implementation of two-dimensional finite-extent impulse response (FIR) filters. The direct, frequency-domain, and block convolution implementations of these filters are discussed, and design algorithms, including the window method, optimal methods, and the transformation method, are presented.

In Chapters 4 and 5, we examine infinite-extent impulse response (IIR) filters which can be represented by two-dimensional constant-coefficient difference equations. Chapter 4 lays the groundwork by introducing the concepts of the two-dimensional difference equation, the z-transform, stability, and the complex cepstrum. Chapter 5 follows with a discussion of implementation strategies and design techniques for two-dimensional IIR filters, including stabilization techniques.

Chapter 6 discusses the use of multidimensional signal processing in the context of processing signals received by an array of sensors. This broad application area is used as a vehicle for introducing the concepts of beamforming and power spectrum estimation. Beamforming represents a linear filtering approach to the problem of determining the strength and direction of propagating energy, while power spectrum estimation represents a modeling and parameter estimation approach to the same problem. Modern spectrum estimation techniques such as high-resolution, all-pole, and maximum entropy methods are discussed together with the more classical techniques. The significant theoretical differences between the one-dimensional and two-dimensional cases are brought out.

In Chapter 7, we discuss inverse problems in which one tries to deduce or reconstruct a signal from limited measurements and *a priori* information. Three examples are examined: deconvolution with constraints, seismic-wave migration, and signal reconstruction from projections.

In writing this text, we have tried to include topics and examples which illustrate fundamental principles of multidimensional signal processing. (Difficult or advanced material is presented in sections flagged by an asterisk.) Since we did not intend to compile an encyclopedic compendium of this still-evolving field, our book is necessarily incomplete. Some may view this as a deficiency, but our goal was to give a broad rather than deep view in the hopes of seducing students into pursuing research projects in this field. We have tried to compensate for any lack of depth by including important references to give readers an entrée into the technical literature.

We have enjoyed the encouragement and patience of many people during the five years it took us to conceive, outline, write, rewrite, and polish this manuscript. Our families were very supportive and patient; time working on the book was generally time spent away from them. Al Oppenheim has had a significant influence on this book in his role as series editor. More importantly, he and Ron Schafer have had a significant influence on the authors in their roles as teachers, supervisors, sounding boards, colleagues, models, and friends.

Over the course of our careers, we have enjoyed the opportunity of interacting with several exceptional colleagues who have made important technical contributions to the field of multidimensional digital signal processing and who have stimulated our own thinking and research in this field. In some cases, their technical contributions are presented explicitly in our book; in other cases, their influence has been more subtle. Among these colleagues, whom we also count as friends, are Professors Demetrius Paris and Monson Hayes, Dr. Mark Richards and Ms. Theresa Speake of Georgia Tech, Dr. Gary Shaw, Dr. Thomas Quatieri, and Dr. Stephen Pohlig of Lincoln Laboratory, Dr. James McClellan of Schlumberger, Professor Jae Lim of M.I.T., and Professor Don Johnson of Rice University. We also owe a measure of gratitude to our respective institutions, the Lincoln Laboratory, M.I.T., and the Georgia Institute of Technology, for providing the intellectual environment which encourages the pursuit of excellence in signal processing, as well as in other areas of engineering and science.

DAN E. DUDGEON
RUSSELL M. MERSEREAU

CONTENTS

INTRODUCTION

One of the by-products of the computer revolution has been the emergence of completely new fields of study. Each year, as integrated circuits have become faster, cheaper, and more compact, it has become possible to find feasible solutions to problems of ever-increasing complexity. Because it demands massive amounts of digital storage and comparable quantities of numerical computation, multidimensional digital signal processing is a problem area which has only recently begun to emerge. Despite this fact, it has already provided the solutions to important problems ranging from computer-aided tomography (CAT), a technique for combining x-ray projections from different orientations to create a three-dimensional reconstruction of a portion of the human body, to the design of passive sonar arrays and the monitoring of the earth's resources by satellite. In addition to its many glamorous and humble applications, however, multidimensional digital signal processing also possesses a firm mathematical foundation, which allows us not only to understand what has already been accomplished, but also to explore rationally new problem areas and solution methods as they arise.

Simply stated, a signal is any medium for conveying information, and signal processing is concerned with the extraction of that information. Thus ensembles of time-varying voltages, the density of silver grains on a photographic emulsion, or lists of numbers in the memory of a computer all represent examples of signals. A typical signal processing task involves the transfer of information from one signal to another. A photograph, for example, might be scanned, sampled, and stored in the memory of a computer. In this case, the information is transferred from a variable silver density, to a beam of visible light, to an electrical waveform, and finally to a sequence of numbers, which, in turn, are represented by an arrangement of magnetic domains on a computer disk. The CAT scanner is a more complex example; information about the structure of an unknown object is first transferred to a series of electromagnetic waves, which are then sampled to produce an array of numbers, which, in turn, are processed by a computational algorithm and finally displayed on the phosphor of a cathode ray tube (CRT) screen or on photographic film. The digital processing which is done cannot add to the information, but it can rearrange it so that a human observer can more readily interpret it; instead of looking at multiple shadows the observer is able to look at a cross-sectional view.

Whatever their form, signals are of interest only because of the information they contain. At the risk of overgeneralizing we might say that signal processing is concerned with two basic tasks—information rearrangement and information reduction. We have already seen two examples of information rearrangement—computer-aided tomography and image scanning. To those we could easily add other examples: image enhancement, image deblurring, spectral analysis, and so on. Information reduction is concerned with the removal of extraneous information. Someone observing radar returns is generally interested in only a few bits of information, specifically, the answer to such questions as: Is anything there? If so, what? Friend or foe? How fast is it going, and where is it headed? However, the receiver is also giving the observer information about the weather, chaff, birds, nearby buildings, noise in the receiver, and so on. The observer must separate the relevant from the irrelevant, and signal processing can help. Other examples of information-lossy signal processing operations include noise removal, parameter estimation, and feature extraction.

Digital signal processing is concerned with the processing of signals which can be represented as sequences of numbers and *multidimensional digital signal processing* is, more specifically, concerned with the processing of signals which can be represented as multidimensional arrays, such as sampled images or sampled time waveforms which are received simultaneously from several sensors. The restriction to digital signals permits processing with digital hardware, and it permits signal processing operators to be specified as algorithms or procedures.

The motivations for looking at digital methods hardly need to be enumerated. Digital methods are simultaneously powerful and flexible. Digital systems can be designed to be adaptive and they can be made to be easily reconfigured. Digital algorithms can be readily transported from the equipment of one manufacturer to another or they can be implemented with special-purpose digital hardware. They can be used equally well to process signals that originated as time functions or as spatial functions and they interface naturally with logical operators such as pattern classifiers. Digital signals can be stored indefinitely without error. For many applications, digital methods may be cheaper than the alternatives, and for others there may simply be no alternatives.

Is the processing of multidimensional signals that different from the processing of one-dimensional ones? At an abstract level, the answer is no. Many operations that we might want to perform on multidimensional sequences are also performed on one-dimensional ones—sampling, filtering, and transform computation, for example. At a closer level, however, we would be forced to say that multidimensional signal processing can be quite different. This is due to three factors: (1) two-dimensional problems generally involve considerably more data than one-dimensional ones; (2) the mathematics for handling multidimensional systems is less complete than the mathematics for handling one-dimensional systems; and (3) multidimensional systems have many more degrees of freedom, which give a system designer a flexibility not encountered in the one-dimensional case. Thus, while all recursive digital filters are implemented using difference equations, in the one-dimensional case these difference equations are totally ordered, whereas in the multidimensional case they are only partially ordered. Flexibility can be exploited. In the one-dimensional case, the discrete Fourier transform (DFT) can be evaluated using the fast Fourier transform (FFT) algorithm, whereas in the multidimensional case, there are a host of DFTs and each can be evaluated using a host of FFT algorithms. In the one-dimensional case, we can adjust the rate at which a bandlimited signal is sampled; in the multidimensional case, we can adjust not only the rate, but also the geometric arrangement of the samples. On the other hand, multidimensional polynomials cannot be factored, whereas one-dimensional ones can. Thus, in the multidimensional case, we cannot talk about isolated poles, zeros, and roots. Multidimensional digital signal processing can be quite different from one-dimensional digital signal processing.

In the early 1960s, many of the methods of one-dimensional digital signal processing were developed with the intention of using the digital systems to simulate analog ones. As a result, much of discrete systems theory was modeled after analog systems theory. In time, it became recognized that, while digital systems could simulate analog systems very well, they could also do much more. With this awareness and a strong push from the technology of digital hardware, the field has blossomed and many of the methods in common use today have no analog equivalents. The same trend can be observed in the development of multidimensional digital signal processing. Since there is no continuous-time or analog two-dimensional systems theory to imitate, early multidimensional systems were based on one-dimensional systems. In the late 1960s, most two-dimensional signal processing was performed using separable two-dimensional systems, which are little more than one-dimensional systems applied to two-dimensional data. In time, uniquely multidimensional algorithms were developed which correspond to logical extrapolations of one-dimensional algorithms. This period was one of frustration. The volume of data demanded by many two-dimensional applications and the absence of a factorization theorem for two-dimensional polynomials meant that many one-dimensional methods did not generalize well. Chronologically, we are now at the dawn of the age of awareness. The computer industry, by making components smaller and cheaper, has helped to solve the data volume problem and we are recognizing that, although we will always have the problem of limited mathematics, multidimensional systems also give us new freedoms. These combine to make the field both challenging and fun.

In this book we summarize many of the advances that have taken place in this exciting and rapidly growing field. The area is one that has evolved with technology. Although we do describe many applications of our material, we have tried not to make it too technology dependent, lest it become technologically obsolete. Rather, we emphasize fundamental concepts so that the reader will not only understand what has been done but will also be able to extend those methods to new applications.

To accomplish all of this, it is necessary to assume some background on the part of the reader. Specifically, we assume that the reader is familiar with one-dimensional linear systems theory and has a basic understanding of one-dimensional digital signal processing (at the level of Oppenheim and Schafer [1], Chaps. 1–6).

In this book our interest is in the processing of all signals of dimensionality greater than or equal to 2. Whereas there is a substantial difference between the theories for the processing of one- and two-dimensional signals, there seems to be little difference between the two-dimensional and higher-dimensional cases, except for the issue of computational complexity. To avoid cluttering up the discussions, equations, and figures of the book, we therefore state the majority of our results only for the two-dimensional case, which is the most prevalent one in applications. In most cases, the generalizations are straightforward, and when they are not, they will be explicitly given. In a similar spirit, we do not belabor results that are obvious generalizations of the one-dimensional case.

We hope the reader will find what we found when we first became involved in the area of multidimensional digital signal processing. It is an area to which a great deal of intuitition may be carried over from the one-dimensional causal world, and yet there are many places where the final form of a result is unexpected and its implications are surprising and counterintuitive. The way in which some one-dimensional results generalize to several dimensions can give the reader new insights into the structure of multidimensional as well as one-dimensional signal processing operations.

REFERENCE

1. Alan V. Oppenheim and Ronald Schafer, *Digital Signal Processing* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1975).

B1

1

MULTIDIMENSIONAL SIGNALS AND SYSTEMS

A multidimensional signal can be modeled as a function of M independent variables, where $M \geq 2$. These signals may be classified as continuous, discrete, or mixed. A *continuous* signal can be modeled as a function of independent variables which range over a continuum of values. For example, the intensity $I(x, y)$ of a photographic image is a two-dimensional continuous signal. A *discrete* signal, on the other hand, can be modeled as a function defined only on a set of points, such as the set of integers. A *mixed* signal is a multidimensional signal that is modeled as a function of some continuous variables and some discrete ones. For example, an ensemble of time waveforms recorded from an array of electrical transducers is a mixed signal. The ensemble can be modeled with one continuous variable, time, and one or more discrete variables to index the transducers.

In this chapter we are concerned primarily with multidimensional discrete signals and the systems that can operate on them. Most of the properties of signals and systems that we will discuss are simple extensions of the properties of one-dimensional discrete signals and systems and therefore, most of our discussions will be brief. The reader who desires further details is referred to one of several excellent textbooks that cover the one-dimensional case [1–3]. It will become apparent, however, that many familiar one-dimensional procedures do not readily generalize to the multidimensional case and that many important issues associated with multidimen-