Advanced Digital Signal Processing and Noise Reduction

Third Edition

Saeed V. Vaseghi
Professor of Communications and Signal Processing
Department of Electronics and Computer Engineering
Brunel University, UK

John Wiley & Sons, Ltd

Advanced Digital Signal Processing and Noise Reduction

Advanced Digital Signal Processing and Noise Reduction

Third Edition

Saeed V. Vaseghi
Professor of Communications and Signal Processing
Department of Electronics and Computer Engineering
Brunel University, UK

John Wiley & Sons, Ltd

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging in Publication Data

Vaseghi, Saeed V.
Advanced digital signal processing and noise reduction / Saeed V. Vaseghi. — 3rd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-470-09494-X
1. Signal processing. 2. Electronic noise. 3. Digital filters (Mathematics) I. Title.
TK5102.9.V37 2005
621.382'2—dc22

2005018514

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN-13 978-0-470-09494-5 (HB)
ISBN-10 0-470-09494-X (HB)

Typeset in 10/12pt Times by Integra Software Services Pvt. Ltd, Pondicherry, India.
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire.
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

To my Luke

I wish to thank Esfandiar Zavarehei and Wendy Pillar for proof reading this edition and for many excellent suggestions. Thanks also to Ben Milner, Qin Yan, Dimitrios Rentzo, Charles Ho and Aimin Chen.

Contents

Preface

The applications of DSP are numerous and include multimedia technology, audio signal processing, video signal processing, cellular mobile communication, adaptive network management, radar systems, pattern analysis, pattern recognition, medical signal processing, financial data forecasting, artificial intelligence, decision making systems, control systems and information search engines.

The theory and application of signal processing is concerned with the identification, modelling and utilisation of patterns and structures in a signal process. The observation signals are often distorted, incomplete and noisy. Hence, noise reduction and the removal of channel distortion and interference are important parts of a signal processing system.

Since the publication of the first edition of this book in 1996, digital signal processing (DSP) in general and noise reduction in particular, have become even more central to the research and development of efficient, adaptive and intelligent mobile communication and information processing systems. The third edition of this book has been revised extensively and improved in several ways to take account of the recent advances in theory and application of digital signal processing. The existing chapters have been updated with new materials added. Two new chapters have been introduced; one for speech enhancement in mobile noisy conditions and the other for modelling and combating noise and fading in wireless communication systems.

The aim of this book is to provide a coherent and structured presentation of the theory and applications of statistical signal processing and noise reduction methods and is organised in 17 chapters.

Chapter 1 begins with an introduction to signal processing, and provides a brief review of signal processing methodologies and applications. The basic operations of sampling and quantisation are reviewed in this chapter.

Chapter 2 provides an introduction to noise and distortion. Several different types of noise, including thermal noise, shot noise, acoustic noise, electromagnetic noise and channel distortions, are considered. The chapter concludes with an introduction to the modelling of noise processes.

Chapter 3 provides an introduction to the theory and applications of probability models and stochastic signal processing. The chapter begins with an introduction to random signals, stochastic processes, probabilistic models and statistical measures. The concepts of stationary, nonstationary and ergodic processes are introduced in this chapter, and some important classes of random processes, such as Gaussian, mixture Gaussian, Markov chains and Poisson processes, are considered. The effects of transformation of a signal on its statistical distribution are considered.

Chapter 4 is on Bayesian estimation and classification. In this chapter the estimation problem is formulated within the general framework of Bayesian inference. The chapter includes Bayesian theory, classical estimators, the estimate–maximise method, the Cramer–Rao bound on the minimum–variance estimate, Bayesian classification, and the modelling of the space of a random signal. This chapter provides a number of examples on Bayesian estimation of signals observed in noise.

Chapter 5 considers hidden Markov models (HMMs) for nonstationary signals. The chapter begins with an introduction to the modelling of nonstationary signals and then concentrates on the theory and applications of hidden Markov models. The hidden Markov model is introduced as a Bayesian model, and methods of training HMMs and using them for decoding and classification are considered. The chapter also includes the application of HMMs in noise reduction.

Chapter 6 considers Wiener filters. The least square error filter is formulated first through minimisation of the expectation of the squared error function over the space of the error signal. Then a block-signal formulation of Wiener filters and a vector space interpretation of Wiener filters are considered. The frequency response of the Wiener filter is derived through minimisation of mean square error in the frequency domain. Some applications of the Wiener filter are considered, and a case study of the Wiener filter for removal of additive noise provides useful insight into the operation of the filter.

Chapter 7 considers adaptive filters. The chapter begins with the state-space equation for Kalman filters. The optimal filter coefficients are derived using the principle of orthogonality of the innovation signal. The recursive least square (RLS) filter, which is an exact sample-adaptive implementation of the Wiener filter, is derived in this chapter. Then the steepest-descent search method for the optimal filter is introduced. The chapter concludes with a study of the LMS adaptive filters.

Chapter 8 considers linear prediction and sub-band linear prediction models. Forward prediction, backward prediction and lattice predictors are studied. This chapter introduces a modified predictor for the modelling of the short-term and the pitch period correlation structures. A maximum *a posteriori* (MAP) estimate of a predictor model that includes the prior probability density function of the predictor is introduced. This chapter concludes with the application of linear prediction in signal restoration.

Chapter 9 considers frequency analysis and power spectrum estimation. The chapter begins with an introduction to the Fourier transform, and the role of the power spectrum in identification of patterns and structures in a signal process. The chapter considers nonparametric spectral estimation, model-based spectral estimation, the maximum entropy method, and high-resolution spectral estimation based on eigenanalysis.

Chapter 10 considers interpolation of a sequence of unknown samples. This chapter begins with a study of the ideal interpolation of a band-limited signal, a simple model for the effects of a number of missing samples, and the factors that affect interpolation. Interpolators are divided into two categories: polynomial and statistical interpolators. A general form of polynomial interpolation as well as its special forms (Lagrange, Newton, Hermite and cubic spline interpolators) is considered. Statistical interpolators in this chapter include maximum *a posteriori* interpolation, least square error interpolation based on an autoregressive model, time–frequency interpolation, and interpolation through the search of an adaptive codebook for the best signal.

Chapter 11 considers spectral subtraction. A general form of spectral subtraction is formulated and the processing distortions that result from spectral subtraction are considered. The effects of processing distortions on the distribution of a signal are illustrated. The chapter considers methods for removal of the distortions and also nonlinear methods of spectral subtraction. This chapter concludes with an implementation of spectral subtraction for signal restoration.

Chapters 12 and 13 cover the modelling, detection and removal of impulsive noise and transient noise pulses. In Chapter 12, impulsive noise is modelled as a binary-state nonstationary process and several stochastic models for impulsive noise are considered. For removal of impulsive noise, median filters and a method based on a linear prediction model of the signal process are considered. The materials in Chapter 13 closely follow Chapter 12. In Chapter 13, a template-based method, an HMM-based method and an AR model-based method for removal of transient noise are considered.

Chapter 14 covers echo cancellation. The chapter begins with an introduction to telephone line echoes, and considers line echo suppression and adaptive line echo cancellation. Then the problem of acoustic echoes and acoustic coupling between loudspeaker and microphone systems is considered. The chapter concludes with a study of a sub-band echo cancellation system.

Chapter 15 covers blind deconvolution and channel equalisation. This chapter begins with an introduction to channel distortion models and the ideal channel equaliser. Then the Wiener equaliser, blind equalisation using the channel input power spectrum, blind deconvolution based on linear predictive models, Bayesian channel equalisation and blind equalisation for digital communication channels are considered. The chapter concludes with equalisation of maximum phase channels using higher-order statistics.

Chapter 16 covers speech enhancement methods. Speech enhancement in noisy environments improves the quality and intelligibility of speech for human communication and increases the accuracy of automatic speech recognition systems. Noise reduction systems are increasingly important in a range of applications such as mobile phones, hands-free phones, teleconferencing systems and in-car cabin communication systems. This chapter provides an overview of the main methods for single-input and multiple-input speech enhancement in noise.

Chapter 17 covers the issue of noise in wireless communication. Noise, fading and limited radio bandwidth are the main factors that constrain the capacity and the speed of communication on wireless channels. Research and development of communications systems aim to increase the spectral efficiency, defined as the data bits per second per Hertz bandwidth of a communication channel. For improved efficiency, modern mobile communications systems rely on signal processing methods at almost every stage from source coding to the allocation of time bandwidth and space resources. In this chapter we consider how communications signal processing methods are employed for improving the speed and capacity of communications systems.

As an additional resource, this book is supported by a companion website on which lecturers and instructors can find electronic versions of the figures. Please go to ftp://ftp.wiley.co.uk/pub/books/vaseghi3e.

Saeed V. Vaseghi

Symbols

$\boldsymbol{A}$	Matrix of predictor coefficients
a_k	Linear predictor coefficients
$\boldsymbol{a}$	Linear predictor coefficients vector
a_{ij}	Probability of transition from state i to state j in a Markov model
$\alpha_i(t)$	Forward probability in an HMM
$b(m)$	Backward prediction error
$b(m)$	Binary state signal
$\beta_i(t)$	Backward probability in an HMM
$c_{xx}(m)$	Covariance of signal $x(m)$
$c_{XX}(k_1, k_2, \cdots, k_N)$	kth-order cumulant of $x(m)$
$C_{XX}(\omega_1, \omega_2, \cdots, \omega_{K-1})$	kth-order cumulant spectra of $x(m)$
$\boldsymbol{D}$	Diagonal matrix
$e(m)$	Estimation error
$\mathcal{E}[x]$	Expectation of x
f	Frequency variable
Fs	Sampling frequency
$f_X(\boldsymbol{x})$	Probability density function for process $\boldsymbol{X}$
$f_{X,Y}(\boldsymbol{x}, \boldsymbol{y})$	Joint probability density function of $\boldsymbol{X}$ and $\boldsymbol{Y}$
$f_{X\|Y}(\boldsymbol{x}\|\boldsymbol{y})$	Probability density function of $\boldsymbol{X}$ conditioned on $\boldsymbol{Y}$
$f_{X;\Theta}(x; \boldsymbol{\theta})$	Probability density function of $\boldsymbol{X}$ with $\boldsymbol{\theta}$ as a parameter
$f_{X\|S,\mathcal{M}}(\boldsymbol{x}\|\boldsymbol{s}, \mathcal{M})$	Probability density function of $\boldsymbol{X}$ given a state sequence $\boldsymbol{s}$ of an HMM $\mathcal{M}$ of the process $\boldsymbol{X}$
$\Phi(m, m-1)$	State transition matrix in Kalman filter
G	Filter gain factor
$\boldsymbol{h}$	Filter coefficient vector, channel response
$\boldsymbol{h}_{\max}$	Maximum-phase channel response
$\boldsymbol{h}_{\min}$	Minimum-phase channel response
$\boldsymbol{h}^{\text{inv}}$	Inverse channel response
$H(f)$	Channel frequency response
$H^{\text{inv}}(f)$	Inverse channel frequency response

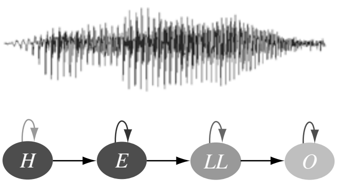
$\boldsymbol{H}$	Observation matrix, distortion matrix
$\boldsymbol{I}$	Identity matrix
$\boldsymbol{J}$	Fisher's information matrix
$\lvert\boldsymbol{J}\rvert$	Jacobian of a transformation
$\boldsymbol{K}(m)$	Kalman gain matrix
λ	Eigenvalue
Λ	Diagonal matrix of eigenvalues
m	Discrete time index
m_k	kth-order moment
$\mathcal{M}$	A model, e.g. an HMM
$\boldsymbol{\mu}$	Adaptation convergence factor
$\boldsymbol{\mu}_x$	Expected mean of vector $\boldsymbol{x}$
$n(m)$	Noise
$\boldsymbol{n}(m)$	A noise vector of N samples
$n_{\mathrm{i}}(m)$	Impulsive noise
$N(f)$	Noise spectrum
$N^*(f)$	Complex conjugate of $N(f)$
$\overline{N(f)}$	Time-averaged noise spectrum
$\mathcal{N}(\boldsymbol{x}, \mu_{\boldsymbol{xx}}, \boldsymbol{\Sigma}_{\boldsymbol{xx}})$	A Gaussian pdf with mean vector $\boldsymbol{\mu}_{\boldsymbol{xx}}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{xx}}$
$O(\cdot)$	In the order of $(\cdot)$
P	Filter order (length)
$P_X(\boldsymbol{x}_i)$	Probability mass function of $\boldsymbol{x}_i$
$P_{X,Y}(\boldsymbol{x}_i, \boldsymbol{y}_j)$	Joint probability mass function of $\boldsymbol{x}_i$ and $\boldsymbol{y}_j$
$P_{X\mid Y}\left(\boldsymbol{x}_i \middle\vert \boldsymbol{y}_j\right)$	Conditional probability mass function of $\boldsymbol{x}_i$ given $\boldsymbol{y}_j$
$P_{\mathrm{NN}}(f)$	Power spectrum of noise $n(m)$
$P_{XX}(f)$	Power spectrum of the signal $x(m)$
$P_{XY}(f)$	Cross-power spectrum of signals $x(m)$ and $y(m)$
$\boldsymbol{\theta}$	Parameter vector
$\hat{\boldsymbol{\theta}}$	Estimate of the parameter vector θ
r_k	Reflection coefficients
$r_{xx}(m)$	Autocorrelation function
$\boldsymbol{r}_{\boldsymbol{xx}}(m)$	Autocorrelation vector
$\boldsymbol{R}_{\boldsymbol{xx}}$	Autocorrelation matrix of signal $\boldsymbol{x}(m)$
$\boldsymbol{R}_{\boldsymbol{xy}}$	Cross-correlation matrix
$\boldsymbol{s}$	State sequence
$\boldsymbol{s}^{\mathrm{ML}}$	Maximum-likelihood state sequence
σ_{n}^2	Variance of noise $n(m)$
$\boldsymbol{\Sigma}_{\mathrm{nn}}$	Covariance matrix of noise $\boldsymbol{n}(m)$
$\boldsymbol{\Sigma}_{\mathrm{xx}}$	Covariance matrix of signal $\boldsymbol{x}(m)$
σ_x^2	Variance of signal $x(m)$
σ_n^2	Variance of noise $n(m)$
$x(m)$	Clean signal
$\hat{x}(m)$	Estimate of clean signal
$\boldsymbol{x}(m)$	Clean signal vector
$X(f)$	Frequency spectrum of signal $x(m)$
$X^*(f)$	Complex conjugate of $X(f)$

$\overline{X(f)}$	Time-averaged frequency spectrum of the signal $x(m)$
$X(f, t)$	Time-frequency spectrum of the signal $x(m)$
$\boldsymbol{X}$	Clean signal matrix
$\boldsymbol{X}^{\mathrm{H}}$	Hermitian transpose of $\boldsymbol{X}$
$y(m)$	Noisy signal
$\boldsymbol{y}(m)$	Noisy signal vector
$\hat{\boldsymbol{y}}(m \mid m-i)$	Prediction of $\boldsymbol{y}(m)$ based on observations up to time m–i
$\boldsymbol{Y}$	Noisy signal matrix
$\boldsymbol{Y}^{\mathrm{H}}$	Hermitian transpose of $\boldsymbol{Y}$
Var	Variance
w_k	Wiener filter coefficients
$\boldsymbol{w}(m)$	Wiener filter coefficients vector
$W(f)$	Wiener filter frequency response
z	z-transform variable

Abbreviations

AR	Autoregressive process
ARMA	Autoregressive moving average process
AWGN	Additive white Gaussian noise
bps	Bits per second
cdf	Cumulative density function
CELP	Code excited linear prediction
dB	Decibels: $10\log_{10}$(power ratio)
DFT	Discrete Fourier transform
DSP	Digital signal processing
EM	Estimate–maximise
ESPIRIT	Estimation of signal parameters via rotational invariance techniques
FFT	Fast Fourier transform
FIR	Finite impulse response
GMM	Gaussian mixture model
GSM	Global system for mobile communications
HMM	Hidden Markov model
Hz	Hertz, unit of frequency in cycles per second
IFFT	Inverse fast Fourier transform
IID	Independent identically distributed
IIR	Infinite impulse response
ISD	Itakura–Saito distance
ISI	Inter symbol interference
LMS	Least mean squared error
LP	Linear prediction model
LPSS	Spectral subtraction based on linear prediction model
LS	Least square
LSAR	Least square AR interpolation
LSE	Least square error
LTI	Linear time invariant
MA	Moving average process
MAP	Maximum *a posteriori* estimate

M-ary	Multilevel signalling
MAVE	Minimum absolute value of error estimate
MIMO	Multiple-input multiple-output
ML	Maximum likelihood estimate
MMSE	Minimum mean squared error estimate
ms	Milliseconds
MUSIC	Multiple signal classification
NLMS	Normalised least mean squared error
pdf	Probability density function
pmf	Probability mass function
psd	Power spectral density
QRD	Orthogonal matrix decomposition
RF	Radio frequency
RLS	Recursive least square
SINR	Signal-to-impulsive noise ratio
SNR	Signal-to-noise ratio
STFT	Short-time Fourier transform
SVD	Singular value decomposition
Var	Variance



1

Introduction

Signal processing provides the basic analysis, modelling and synthesis tools for a diverse area of technological fields, including telecommunication, artificial intelligence, biological computation and system identification. Signal processing is concerned with the modelling, detection, identification and utilisation of patterns and structures in a signal process. Applications of signal processing methods include audio hi-fi, digital TV and radio, cellular mobile phones, voice recognition, vision, radar, sonar, geophysical exploration, medical electronics, bio-signal processing and in general any system that is concerned with the communication or processing and retrieval of information. Signal processing theory plays a central role in the development of digital telecommunication and automation systems, and in the efficient transmission, reception and decoding of information.

This chapter begins with a definition of signals, and a brief introduction to various signal processing methodologies. We consider several key applications of digital signal processing in adaptive noise reduction, channel equalisation, pattern classification/recognition, audio signal coding, signal detection, spatial processing for directional reception of signals, Dolby noise reduction and radar.

1.1 SIGNALS AND INFORMATION

A signal is the variation of a quantity by which information is conveyed regarding the state, the characteristics, the composition, the trajectory, the evolution, the course of action or the

Advanced Digital Signal Processing and Noise Reduction Third Edition Saeed V. Vaseghi

intention of the information source. A signal is a means of conveying information regarding the state(s) of a variable.

The information conveyed in a signal may be used by humans or machines for communication, forecasting, decision-making, control, geophysical exploration, medical diagnosis, forensics, etc. The types of signals that signal processing deals with include textual data, audio, ultrasonic, subsonic, image, electromagnetic, medical, biological, financial and seismic signals.

Figure 1.1 illustrates a communication system composed of an information source, $I(t)$, followed by a system, $T[\cdot]$, for transformation of the information into variation of a signal, $x(t)$, a communication channel, $h[\cdot]$, for propagation of the signal from the transmitter to the receiver, additive channel noise, $n(t)$, and a signal processing unit at the receiver for extraction of the information from the received signal.

In general, there is a mapping operation that maps the output, $I(t)$, of an information source to the signal, $x(t)$, that carries the information; this mapping operator may be denoted as $T[\cdot]$ and expressed as

$$x(t) = T[I(t)] \tag{1.1}$$

The information source $I(t)$ is normally discrete-valued, whereas the signal $x(t)$ that carries the information to a receiver may be continuous or discrete. For example, in multimedia communication the information from a computer, or any other digital communication device, is in the form of a sequence of binary numbers (ones and zeros), which would need to be transformed into voltage or current variations and modulated to the appropriate form for transmission in a communication channel over a physical link.

As a further example, in human speech communication the voice-generating mechanism provides a means for the speaker to map each discrete word into a distinct pattern of modulation of the acoustic vibrations of air that can propagate to the listener. To communicate a word, w, the speaker generates an acoustic signal realisation of the word, $x(t)$; this acoustic signal may be contaminated by ambient noise and/or distorted by a communication channel, or impaired by the speaking abnormalities of the talker, and received as the noisy, distorted and/or incomplete signal $y(t)$, modelled as

$$y(t) = h[x(t)] + n(t) \tag{1.2}$$

In addition to conveying the spoken word, the acoustic speech signal has the capacity to convey information on the prosody (i.e. pitch, intonation and stress patterns in pronunciation) of speech and the speaking characteristics, accent and emotional state of the talker. The listener extracts this information by processing the signal $y(t)$.

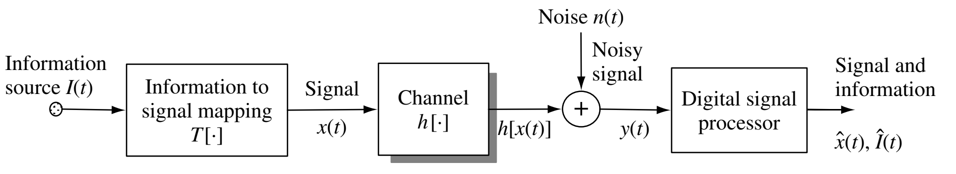


Figure 1.1 Illustration of a communication and signal processing system.

In the past few decades, the theory and applications of digital signal processing have evolved to play a central role in the development of modern telecommunication and information technology systems.

Signal processing methods are central to efficient communication, and to the development of intelligent man–machine interfaces in areas such as speech and visual pattern recognition for multimedia systems. In general, digital signal processing is concerned with two broad areas of information theory:

(1) efficient and reliable coding, transmission, reception, storage and representation of signals in communication systems; and
(2) extraction of information from noisy signals for pattern recognition, detection, forecasting, decision-making, signal enhancement, control, automation, etc.

In the next section we consider four broad approaches to signal processing.

1.2 SIGNAL PROCESSING METHODS

Signal processing methods have evolved in algorithmic complexity, aiming for optimal utilisation of the information in order to achieve the best performance. In general the computational requirement of signal processing methods increases, often exponentially, with the algorithmic complexity. However, the implementation cost of advanced signal processing methods has been offset and made affordable by the consistent trend in recent years of a continuing increase in the performance, coupled with a simultaneous decrease in the cost, of signal processing hardware.

Depending on the method used, digital signal processing algorithms can be categorised into one or a combination of four broad categories. These are transform-based signal processing, model-based signal processing, Bayesian statistical signal processing and neural networks, as illustrated in Figure 1.2. These methods are briefly described below.

1.2.1 TRANSFORM-BASED SIGNAL PROCESSING

The purpose of a transform is to describe a signal or a system in terms of a combination of a set of elementary simple signals (such as sinusoidal signals) that lend themselves to

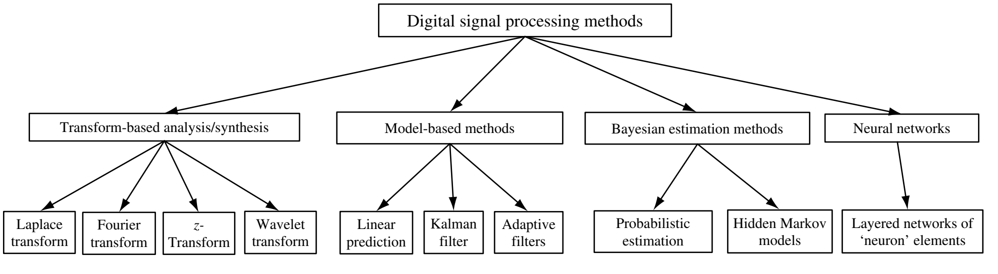


Figure 1.2 A broad categorisation of some of the most commonly used signal processing methods.

relatively easy analysis, interpretation and manipulation. Transform-based signal processing methods include Fourier transform, Laplace transform, z-transform and wavelet transforms. The most widely applied signal transform is the Fourier transform, which is effectively a form of vibration analysis, in that a signal is expressed in terms of a combination of the sinusoidal vibrations that make up the signal. Fourier transform is employed in a wide range of applications, including popular music coders, noise reduction and feature extraction for pattern recognition. The Laplace transform, and its discrete-time version the z-transform, are generalisations of the Fourier transform and describe a signal or a system in terms of a set of sinusoids with exponential amplitude envelopes.

In Fourier, Laplace and z-transform, the different sinusoidal basis functions of the transforms all have the same duration and differ in terms of their frequency of vibrations and amplitude envelopes. In contrast, the wavelets are multi-resolution transforms in which a signal is described in terms of a combination of elementary waves of different durations. The set of basis functions in a wavelet is composed of contractions and dilations of a single elementary wave. This allows non-stationary events of various durations in a signal to be identified and analysed.

1.2.2 MODEL-BASED SIGNAL PROCESSING

Model-based signal processing methods utilise a parametric model of the signal generation process. The parametric model normally describes the predictable structures and the expected patterns in the signal process, and can be used to forecast the future values of a signal from its past trajectory. Model-based methods normally outperform nonparametric methods, since they utilise more information in the form of a model of the signal process. However, they can be sensitive to the deviations of a signal from the class of signals characterised by the model. The most widely used parametric model is the linear prediction model, described in Chapter 8. Linear prediction models have facilitated the development of advanced signal processing methods for a wide range of applications such as low-bit-rate speech coding in cellular mobile telephony, digital video coding, high-resolution spectral analysis, radar signal processing and speech recognition.

1.2.3 BAYESIAN SIGNAL PROCESSING

The fluctuations of a purely random signal, or the distribution of a class of random signals in the signal space, cannot be modelled by a predictive equation, but can be described in terms of the statistical average values, and modelled by a probability distribution function in a multidimensional signal space. For example, as described in Chapter 10, a linear prediction model driven by a random signal can provide a source-filter model of the acoustic realisation of a spoken word. However, the random input signal of the linear prediction model, or the variations in the characteristics of different acoustic realisations of the same word across the speaking population, can only be described in statistical terms and in terms of probability functions.

The Bayesian inference theory provides a generalised framework for statistical processing of random signals, and for formulating and solving estimation and decision-making problems. Chapter 4 describes the Bayesian inference methodology and the estimation of random processes observed in noise.

1.2.4 NEURAL NETWORKS

Neural networks are combinations of relatively simple nonlinear adaptive processing units, arranged to have a structural resemblance to the transmission and processing of signals in biological neurons. In a neural network several layers of parallel processing elements are interconnected by a hierarchically structured connection network. The connection weights are trained to perform a signal processing function such as prediction or classification. Neural networks are particularly useful in nonlinear partitioning of a signal space, in feature extraction and pattern recognition and in decision-making systems. In some hybrid pattern recognition systems neural networks are used to complement Bayesian inference methods. Since the main objective of this book is to provide a coherent presentation of the theory and applications of statistical signal processing, neural networks are not discussed in this book

1.3 APPLICATIONS OF DIGITAL SIGNAL PROCESSING

In recent years, the development and commercial availability of increasingly powerful and affordable digital computers has been accompanied by the development of advanced digital signal processing algorithms for a wide variety of applications such as noise reduction, telecommunications, radar, sonar, video and audio signal processing, pattern recognition, geophysics explorations, data forecasting, and the processing of large databases for the identification, extraction and organisation of unknown underlying structures and patterns. Figure 1.3 shows a broad categorisation of some digital signal processing (DSP) applications. This section provides a review of several key applications of DSP methods.

1.3.1 ADAPTIVE NOISE CANCELLATION

In speech communication from a noisy acoustic environment such as a moving car or train, or over a noisy telephone channel, the speech signal is observed in an additive random noise.

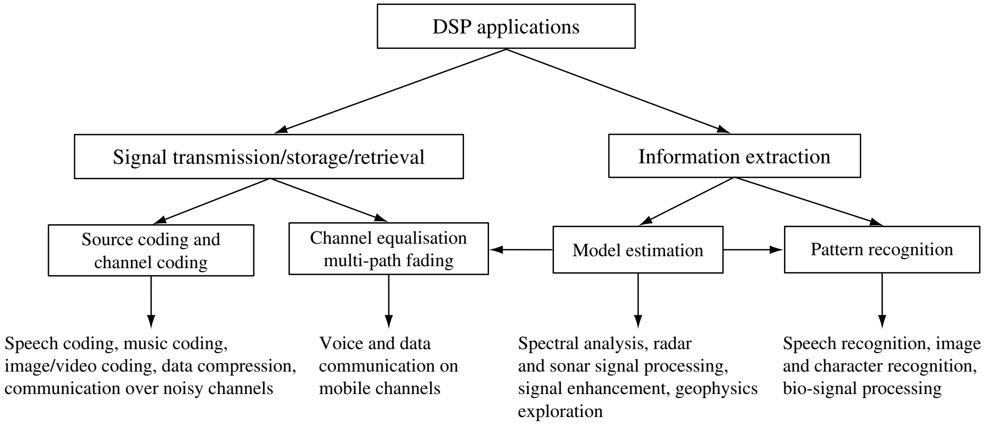


Figure 1.3 A classification of the applications of digital signal processing.

In signal measurement systems the information-bearing signal is often contaminated by noise from its surrounding environment. The noisy observation, $y(m)$, can be modelled as

$$y(m) = x(m) + n(m) \tag{1.3}$$

where $x(m)$ and $n(m)$ are the signal and the noise, and m is the discrete-time index. In some situations, for example when using a mobile telephone in a moving car, or when using a radio communication device in an aircraft cockpit, it may be possible to measure and estimate the instantaneous amplitude of the ambient noise using a directional microphone. The signal, $x(m)$, may then be recovered by subtraction of an estimate of the noise from the noisy signal.

Figure 1.4 shows a two-input adaptive noise cancellation system for enhancement of noisy speech. In this system a directional microphone takes as input the noisy signal $x(m)+n(m)$, and a second directional microphone, positioned some distance away, measures the noise $\alpha n(m+\tau)$. The attenuation factor, α, and the time delay, τ, provide a rather over-simplified model of the effects of propagation of the noise to different positions in the space where the microphones are placed. The noise from the second microphone is processed by an adaptive digital filter to make it equal to the noise contaminating the speech signal, and then subtracted from the noisy signal to cancel out the noise. The adaptive noise canceller is more effective in cancelling out the low-frequency part of the noise, but generally suffers from the nonstationary character of the signals, and from the over-simplified assumption that a linear filter can model the diffusion and propagation of the noise sound in the space.

1.3.2 ADAPTIVE NOISE REDUCTION

In many applications, for example at the receiver of a telecommunication system, there is no access to the instantaneous value of the contaminating noise, and only the noisy signal is available. In such cases the noise cannot be cancelled out, but it may be reduced, in an

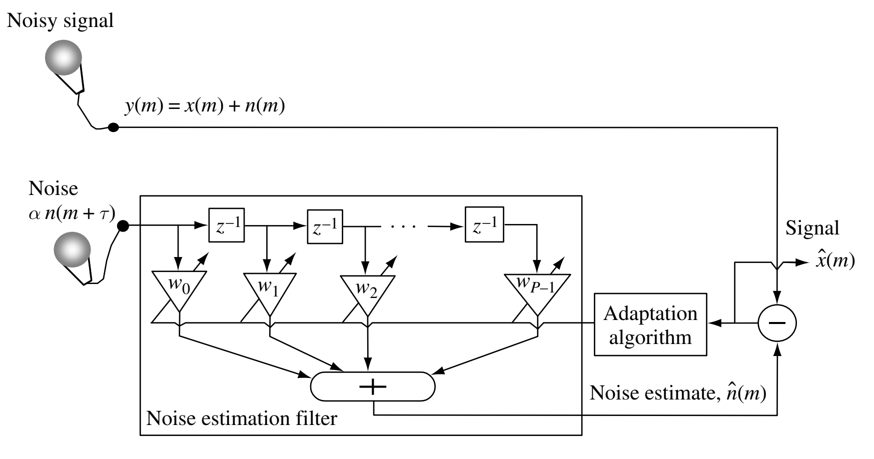


Figure 1.4 Configuration of a two-microphone adaptive noise canceller.

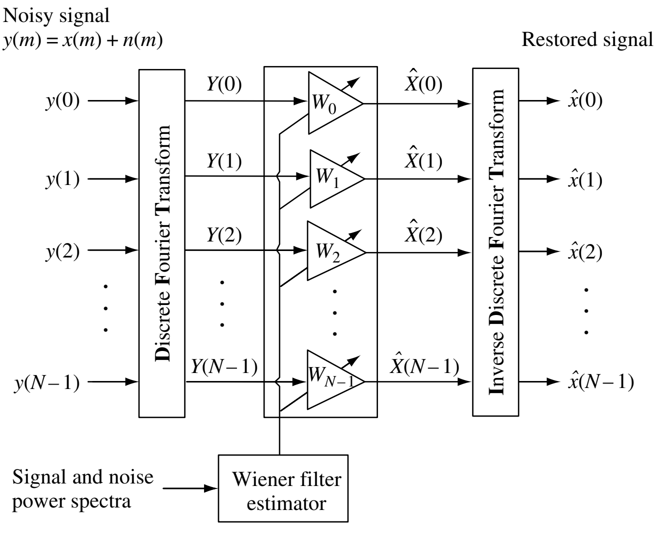


Figure 1.5 A frequency-domain Wiener filter for reducing additive noise.

average sense, using the statistics of the signal and the noise process. Figure 1.5 shows a bank of Wiener filters for reducing additive noise when only the noisy signal is available. The filter bank coefficients attenuate each noisy signal frequency in inverse proportion to the signal-to-noise ratio at that frequency. The Wiener filter bank coefficients, derived in Chapter 6, are calculated from estimates of the power spectra of the signal and the noise processes.

1.3.3 BLIND CHANNEL EQUALISATION

Channel equalisation is the recovery of a signal distorted in transmission through a communication channel with a nonflat magnitude or a nonlinear phase response. When the channel response is unknown, the process of signal recovery is called ‘blind equalisation’. Blind equalisation has a wide range of applications, for example in digital telecommunications for removal of inter-symbol interference due to nonideal channel and multipath propagation, in speech recognition for removal of the effects of the microphones and communication channels, in correction of distorted images, in analysis of seismic data and in de-reverberation of acoustic gramophone recordings.

In practice, blind equalisation is feasible only if some useful statistics of the channel input are available. The success of a blind equalisation method depends on how much is known about the characteristics of the input signal and how useful this knowledge can be in the channel identification and equalisation process. Figure 1.6 illustrates the configuration of a decision-directed equaliser. This blind channel equaliser is composed of two distinct sections: an adaptive equaliser that removes a large part of the channel distortion, followed by a nonlinear decision device for an improved estimate of the channel input. The output of the decision device is the final estimate of the channel input, and it is used as the desired

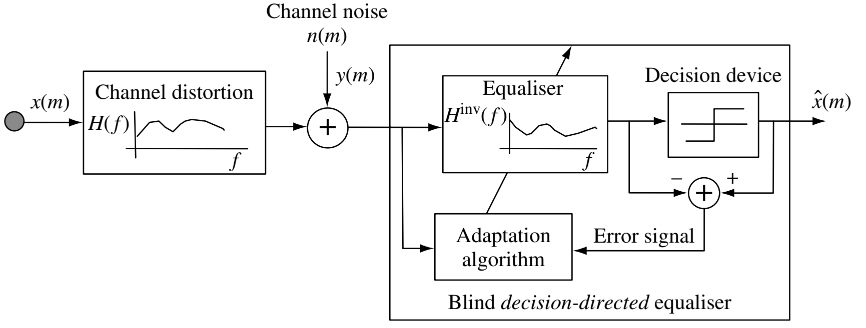


Figure 1.6 Configuration of a decision-directed blind channel equaliser.

signal *to direct* the equaliser adaptation process. Blind equalisation is covered in detail in Chapter 15.

1.3.4 SIGNAL CLASSIFICATION AND PATTERN RECOGNITION

Signal classification is used in detection, pattern recognition and decision-making systems. For example, a simple binary-state classifier can act as the detector of the presence, or the absence, of a known waveform in noise. In signal classification, the aim is to design a minimum-error system for *labelling* a signal with one of a number of likely classes of signal.

To design a classifier, a set of models is trained for the classes of signals that are of interest in the application. The simplest form that the models can assume is a bank, or code book, of waveforms, each representing the prototype for one class of signals. A more complete model for each class of signals takes the form of a probability distribution function. In the classification phase, a signal is labelled with the nearest or the most likely class. For example, in communication of a binary bit stream over a band-pass channel, the binary phase-shift keying (BPSK) scheme signals the bit '1' using the waveform $A_c \sin \omega_c t$ and the bit '0' using $-A_c \sin \omega_c t$.

At the receiver, the decoder has the task of classifying and labelling the received noisy signal as a '1' or a '0'. Figure 1.7 illustrates a correlation receiver for a BPSK signalling

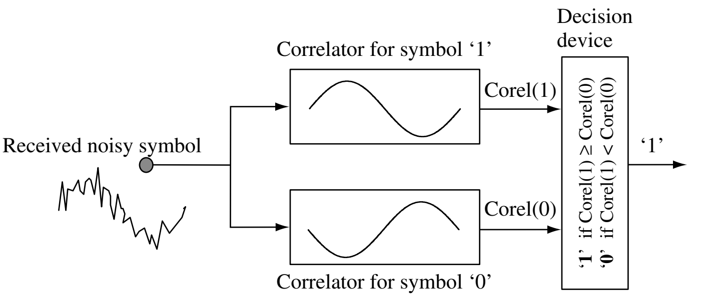


Figure 1.7 A block diagram illustration of the classifier in a binary phase-shift keying demodulation.

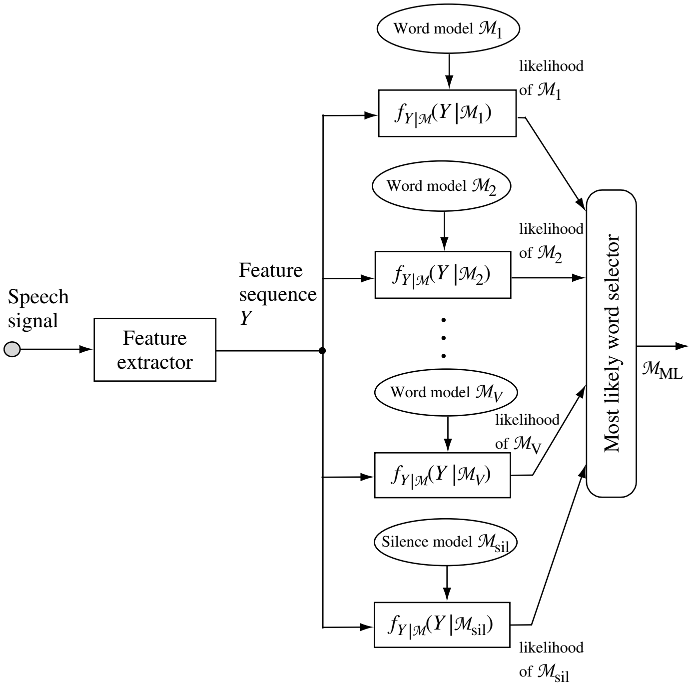


Figure 1.8 Configuration of a speech recognition system; $f(Y|\mathcal{M}_i)$ is the likelihood of the model $\mathcal{M}_i$ given an observation sequence $\boldsymbol{Y}$.

scheme. The receiver has two correlators, each programmed with one of the two symbols representing the binary states for the bit ‘1’ and the bit ‘0’. The decoder correlates the unlabelled input signal with each of the two candidate symbols and selects the candidate that has a higher correlation with the input.

Figure 1.8 illustrates the use of a classifier in a limited-vocabulary, isolated-word speech recognition system. Assume there are V words in the vocabulary. For each word a model is trained, on many different examples of the spoken word, to capture the average characteristics and the statistical variations of the word. The classifier has access to a bank of $V+1$ models, one for each word in the vocabulary and an additional model for the silence periods. In the speech-recognition phase, the task is to decode and label an acoustic speech feature sequence, representing an unlabelled spoken word, as one of the V likely words or silence. For each candidate word the classifier calculates a probability score and selects the word with the highest score.

1.3.5 LINEAR PREDICTION MODELLING OF SPEECH

Linear predictive models are widely used in speech processing applications such as low-bit-rate speech coding in cellular telephony, speech enhancement and speech recognition. Speech is generated by inhaling air into the lungs, and then exhaling it through the vibrating

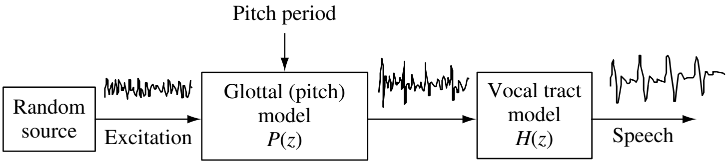


Figure 1.9 Linear predictive model of speech.

glottis cords and the vocal tract. The random, noise-like, air flow from the lungs is spectrally shaped and amplified by the vibrations of the glottal cords and the resonance of the vocal tract. The effect of the vibrations of the glottal cords and the vocal tract is to introduce a measure of correlation and predictability to the random variations of the air from the lungs. Figure 1.9 illustrates a source-filter model for speech production. The source models the lung and emits a random excitation signal which is filtered, first by a pitch filter model of the glottal cords and then by a model of the vocal tract.

The main source of correlation in speech is the vocal tract modelled by a linear predictor. A linear predictor forecasts the amplitude of the signal at time $m, x(m)$, using a linear combination of P previous samples $[x(m-1), \cdots, x(m-P)]$ as

$$\hat{x}(m) = \sum_{k=1}^{P} a_k x(m-k) \tag{1.4}$$

where $\hat{x}(m)$ is the prediction of the signal $x(m)$, and the vector $\boldsymbol{a}^{\mathrm{T}} = [a_1, \ldots, a_P]$ is the coefficients vector of a predictor of order P. The prediction error $e(m)$, i.e. the difference between the actual sample, $x(m)$, and its predicted value, $\hat{x}(m)$, is defined as

$$e(m) = x(m) - \sum_{k=1}^{P} a_k x(m-k) \tag{1.5}$$

The prediction error $e(m)$ may also be interpreted as the random excitation or the so-called innovation content of $x(m)$. From Equation (1.5) a signal generated by a linear predictor can be synthesised as

$$x(m) = \sum_{k=1}^{P} a_k x(m-k) + e(m) \tag{1.6}$$

1.3.6 DIGITAL CODING OF AUDIO SIGNALS

In digital audio, the memory required to record a signal, the bandwidth required for signal transmission and the signal-to-quantisation noise ratio are all directly proportional to the number of bits per sample. The objective in the design of a coder is to achieve high fidelity with as few bits per sample as possible, at an affordable implementation cost. Audio signal coding schemes utilise the statistical structures of the signal and a model of the signal generation, together with information on the psychoacoustics and the masking effects of hearing. In general, there are two main categories of audio coders: model-based coders, used

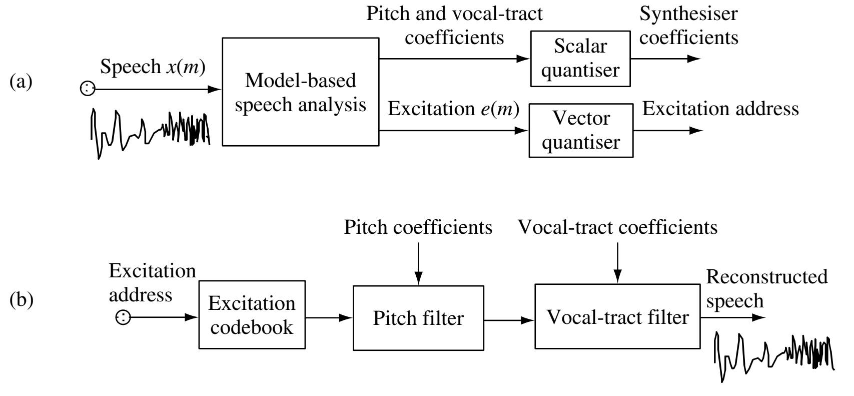


Figure 1.10 Block diagram configuration of a model-based speech (a) coder and (b) decoder.

for low-bit-rate speech coding in applications such as cellular telephony, and transform-based coders used in high-quality coding of speech and digital hi-fi audio.

Figure 1.10 shows a simplified block diagram configuration of a speech coder-decoder of the type used in digital cellular telephones. The speech signal is modelled as the output of a filter excited by a random signal. The random excitation models the air exhaled through the lung, and the filter models the vibrations of the glottal cords and the vocal tract. At the transmitter, speech is segmented into blocks about 30 ms long, during which speech parameters can be assumed to be stationary. Each block of speech samples is analysed to extract and transmit a set of excitation and filter parameters that can be used to synthesise the speech. At the receiver, the model parameters and the excitation are used to reconstruct the speech.

A transform-based coder is shown in Figure 1.11. The aim of transformation is to convert the signal into a form that lends itself to more convenient and useful interpretation and manipulation. In Figure 1.11 the input signal is transformed to the frequency domain using

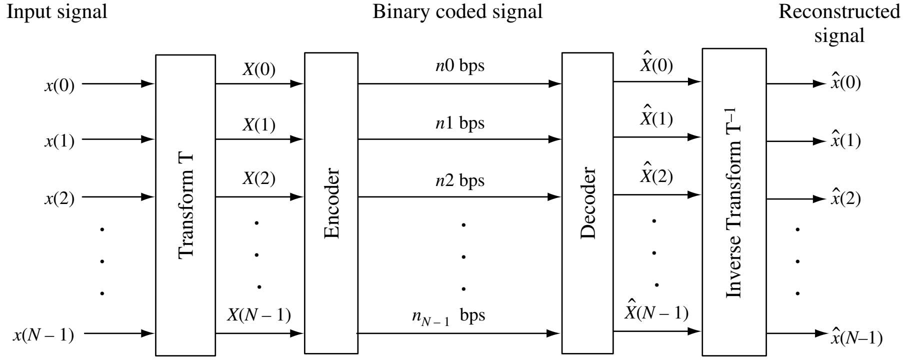


Figure 1.11 Illustration of a transform-based coder.

a filter bank, or a discrete Fourier transform, or a discrete cosine transform. The three main advantages of coding a signal in the frequency domain are:

(1) The frequency spectrum of a signal has a relatively well-defined structure, for example most of the signal power is usually concentrated in the lower regions of the spectrum.
(2) A relatively low-amplitude frequency would be masked in the near vicinity of a large-amplitude frequency and can therefore be coarsely encoded without any audible degradation.
(3) The frequency samples are orthogonal and can be coded independently with different precisions.

The number of bits assigned to each frequency of a signal is a variable that reflects the contribution of that frequency to the reproduction of a perceptually high-quality signal. In an adaptive coder, the allocation of bits to different frequencies is made to vary with the time variations of the power spectrum of the signal.

1.3.7 DETECTION OF SIGNALS IN NOISE

In the detection of signals in noise, the aim is to determine if the observation consists of noise alone, or if it contains a signal. The noisy observation, $y(m)$, can be modelled as

$$y(m) = b(m)x(m) + n(m) \tag{1.7}$$

where $x(m)$ is the signal to be detected, $n(m)$ is the noise and $b(m)$ is a binary-valued state indicator sequence such that $b(m) = 1$ indicates the presence of the signal, $x(m)$, and $b(m) = 0$ indicates that the signal is absent. If the signal, $x(m)$, has a known shape, then a correlator or a matched filter can be used to detect the signal, as shown in Figure 1.12. The impulse response $h(m)$ of the matched filter for detection of a signal, $x(m)$, is the time-reversed version of $x(m)$ given by

$$h(m) = x(N-1-m) \qquad 0 \le m \le N-1 \tag{1.8}$$

where N is the length of $x(m)$. The output of the matched filter is given by

$$z(m) = \sum_{m=0}^{N-1} h(m-k)y(m) \tag{1.9}$$

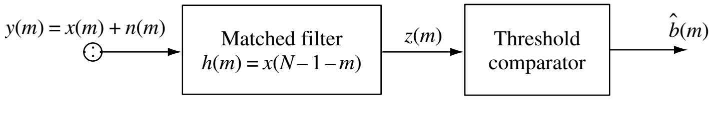


Figure 1.12 Configuration of a matched filter followed by a threshold comparator for detection of signals in noise.