Saeed V. Vaseghi

Multimedia Signal Processing

Theory and Applications in Speech, Music and Communications

Companion Website

WILEY

Multimedia Signal Processing

Theory and Applications in Speech, Music and Communications

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

BICENTENNIAL 1807 WILEY 2007 BICENTENNIAL
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.

Anniversary Logo Design: Richard J. Pacifico

Library of Congress Cataloging in Publication Data

Vaseghi, Saeed V.
Multimedia signal processing: theory and applications in speech, music and communications / Saeed V. Vaseghi.
p. cm. Includes index.
ISBN 978-0-470-06201-2 (alk. paper)
1. Signal processing. 2. Multimedia systems. I. Title.
TK5102.9.V38 2007
621.382′2—dc22

2007014614

British Library Cataloguing in Publication Data

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

ISBN 978-0-470-06201-2 (HB)

Typeset in 9/11pt Times by Integra Software Services Pvt. Ltd, Pondicherry, India
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, England.
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.

Multimedia Signal Processing

To my dears Luke, Jalil, Iran, Geraldine, Simin, Sima, Mona, Malisha

Preface

The applications of digital signal processing (DSP) in information and communication technology are pervasive and include multimedia technology, cellular mobile communication, adaptive network management, radar, pattern recognition, medical signal processing, financial data forecasting, artificial intelligence, decision making systems, control systems and search engines.

The aim of this book is to provide an accessible text to the theory and applications of DSP. This is an ambitious task as signal processing covers such a wide range of topics that it would take several volumes to cover the entire subject area. Nevertheless, I have tried to cover a set of core topics and applications that would enable interested readers to develop a practical understanding of the design and working of modern application of signal processing.

For example, the popular MP3 music coder is based on a combination of a number of signal processing tools including filter banks, discrete Fourier transform, discrete cosine transform, auditory masking thresholds, Huffman code and rate-distortion theory. All these topics are covered in this book as is their combined application in a single system such as MP3.

This book is arranged in three parts and eighteen chapters.

Part I Basic Digital Signal Processing in five chapters introduces the Fourier and z transforms, digital filters, sampling and quantisation.

Chapter 1 begins with an introduction to digital signal processing, and provides a brief review of signal processing methodologies and applications.

Chapter 2 presents the theory and application of Fourier analysis and synthesis of signals. This chapter covers discrete Fourier transform (DFT), fast Fourier transform (FFT) and discrete cosine transform (DCT). Important engineering issues such as time-frequency resolutions, the effect of windowing and spectral leakage are discussed. Several applications of Fourier transform, such as the design of a spectrogram, are presented.

Chapter 3 provides an introduction to z-transform. The relationship between z-transform, Laplace transform and Fourier transform are explored and z-transform is derived from the Laplace transform. The concept of z-transfer function and the associated poles and zeros of the transfer function are introduced.

Chapter 4 introduces the theory and design of digital filters. This chapter starts from a basic explanations of the working of filters and introduces different popular methods for filter design. This chapter also covers the design of quadrature mirror filters which have applications in coding and wavelet analysis.

Chapter 5 introduces the theory and practice of sampling for conversion of analogue signals to discrete-time signals, the process of quantisation of analogue samples of a discrete-time signal to digital values and the process of resampling of a discrete-time signal.

Part II Model-based Signal Processing covers the theory and applications of probability and information models, Bayesian inference, Wiener filters, Kalman filters, adaptive filters, linear prediction models, eigen analysis, principal component analysis (PCA) and independent component analysis (ICA).

Chapter 6 presents probability and information models. This chapter begins with an introduction to random signals, stochastic processes, probabilistic models and statistical measures. The concepts of entropy, stationary, non-stationary and ergodic processes are introduced and some important classes of random processes are considered. The effects of transformation of a signal on its statistical distribution are considered.

Chapter 7 is on Bayesian inference. 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 Cramér–Rao bound on the minimum-variance estimate, Bayesian classification, and the k-means method of modelling of the space of a random signal. This chapter provides a number of examples on Bayesian estimation of signals observed in noise.

Chapter 8 considers Wiener filters. The least square error filter is first formulated 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 the mean square error in the frequency domain. Some applications of the Wiener filter are considered. A case study of the Wiener filter for removal of additive noise provides particularly useful insight into the operation of the filter.

Chapter 9 considers adaptive filters. The chapter begins with the state-space equation for Kalman filters. The recursive least squared (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 least mean squared (LMS) adaptive filters.

Chapter 10 considers linear prediction. Forward prediction, backward prediction, lattice predictors and sub-band 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 11 considers hidden Markov models (HMMs) for non-stationary signals. The chapter begins with an introduction to the modelling of non-stationary 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 12 covers the immensely important and related subjects of eigen analysis, principal component analysis and independent component analysis. Several illustrative example of applications of each method are provided.

Part III Applications of Digital Signal Processing in Speech, Music and Telecommunications covers speech processing, music processing, speech enhancement, echo cancellation and communication signal processing.

Chapter 13 on music signal processing begins with an introduction to musical notes, musical intervals and scales. The physics of some string and pipe instruments namely guitar, violin and trumpet are studied. This chapter includes a study of the anatomy of the ear and the psychoacoustics of hearing. The chapter ends with a study of music coders and MP3.

Chapter 14 on speech processing starts with an introduction to the physiology of speech production. The linear prediction model of speech and the harmonic plus noise model of speech are presented as is a model that combines the two. Speech coding for mobile phones and speech recognition for voice-dialling are considered in this chapter.

Chapter 15 on speech enhancement considers noise reduction, packet loss concealment and bandwidth extension. On noise reduction we consider methods such as spectral subtraction, Wiener filter and Kalman filter. A particular feature of this chapter is the use of a combination of linear prediction model and harmonic plus noise model as a framework for speech enhancement.

Chapter 16 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 problems of acoustic echoes and acoustic coupling between loudspeaker and microphone systems are considered. The chapter concludes with a study of a sub-band echo cancellation system

Chapter 17 is on 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 18 covers wireless communication. Noise, fading and limited radio spectrum are the main factors that constrain the capacity and the speed of communication. For improved efficiency modern mobile communication 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 communication signal processing methods are employed for improving the speed and capacity of communication systems.

The following companion website provides Matlab programs and audio–visual animations related to the subjects covered in this book http://dea.brunel.ac.uk/cmsp/mmsp.

Saeed V. Vaseghi

Acknowledgement

I wish to thank Saeed Ayat for his technical proofreading of this book and his suggestions and contributions. Many thanks to the publishing team at John Wiley, Sarah Hinton, Mark Hammond, Jen Beal, Brett Wells, Katharine Unwin and Lyn Imeson for all their support and contributions. Thanks also to Ales Prochazka, Helena Jetelova, Esi Zavarehei, Emir Tourajlic, Ben Milner, Qin Yan, Dimitrios Rentzos, Charles Ho, Ali Ghorshi, Phuay Low and Aimin Chen.

Symbols

$\boldsymbol{A}$	Matrix of predictor coefficients
a_k	Linear predictor coefficients
$\boldsymbol{a}$	Linear predictor coefficients vector
$\boldsymbol{a}^{inv}$	Inverse 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, Binary state signal
$\beta_i(t)$	Backward probability in an HMM
$\boldsymbol{C}_{xx}$	Covariance matrix of $\boldsymbol{x}$
$c_{xx}(m)$	autocovariance of signal at lag m
$c_{XX}(k_1, k_2, \cdots, k_N)$	k^{th} order cumulant of $x(m)$
$C_{XX}(\omega_1, \omega_2, \cdots, \omega_{k-1})$	k^{th} order cumulant spectra of $x(m)$
$\boldsymbol{D}$	Diagonal matrix
Δf	Frequency resolution
$\delta(t)$	Dirac delta function
$e(m)$	Estimation error or prediction error
$\mathcal{E}[x]$	Expectation of x
f	Frequency variable
f_c	Filter cutoff frequency
F_0	Fundamental frequency
F_s	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 $\boldsymbol{Y}$
$f_{X;\Theta}(\boldsymbol{x};\theta)$	Probability density function of $\boldsymbol{X}$ with θ 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
$\boldsymbol{h}_{\text{max}}$	Maximum-phase channel response
$\boldsymbol{h}_{\text{min}}$	Minimum-phase channel response
$\boldsymbol{h}^{\text{inv}}$	Inverse channel response

$H(f)$	Channel frequency response
$H^{\text{inv}}(f)$	Inverse channel frequency response
$H(z)$	z-transfer function
$\boldsymbol{H}$	Observation matrix, Distortion matrix
$\boldsymbol{I}$	Identity matrix
$\boldsymbol{J}$	Fisher's information matrix
$\lvert\boldsymbol{J}\rvert$	Jacobian of a transformation
JND	Just noticeable distortion level
$\boldsymbol{K}(m)$	Kalman gain matrix
$k(x)$	Kurtosis
λ	Eigenvalue
Λ	Diagonal matrix of eigenvalues
m	Discrete time index
m_k	k^{th} order moment
$\mathcal{M}$	A model, e.g. an HMM
μ	Adaptation step size
μ_x	Expected mean of vector $\boldsymbol{x}$
$n(m)$	Noise
$\boldsymbol{n}(m)$	A noise vector of N samples
$n_i(m)$	Impulsive noise
$N(f)$	Noise spectrum
$N^*(f)$	Complex conjugate of $N(f)$
$\overline{N(f)}$	Time-averaged noise spectrum
$\boldsymbol{N}(\boldsymbol{x}, \mu_{\boldsymbol{xx}}, \Sigma_{\boldsymbol{xx}})$	A Gaussian pdf with mean vector $\mu_{\boldsymbol{xx}}$ and covariance matrix $\Sigma_{\boldsymbol{xx}}$
$O(\cdot)$	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 $\boldsymbol{y}_j$
$P_{X\mid Y}(\boldsymbol{x}_i \mid \boldsymbol{y}_j)$	Conditional probability mass function of $\boldsymbol{x}_i$ given $\boldsymbol{y}_j$
$P_{NN}(f)$	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 $y(m)$
$Q(x_1, x_2, x_3, \ldots)$	Cumulant
θ	Parameter vector
$\hat{\theta}$	Estimate of the parameter vector θ
r_k	Reflection coefficients
$r_{xx}(m)$	Autocorrelation function
$\boldsymbol{r}_{xx}(m)$	Autocorrelation vector
$\boldsymbol{R}_{\boldsymbol{xx}}$	Autocorrelation matrix of signal $\boldsymbol{x}(m)$
$\boldsymbol{R}_{\boldsymbol{xy}}$	Cross-correlation matrix
T_s	Sampling period
$\boldsymbol{s}$	State sequence
$\boldsymbol{s}^{ML}$	Maximum-likelihood state sequence
σ_n^2	Variance of noise $n(m)$
$\Sigma_{\boldsymbol{nn}}$	Covariance matrix of noise $\boldsymbol{n}(m)$
$\Sigma_{\boldsymbol{xx}}$	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
$\hat{\boldsymbol{y}}(m\,\|\,m-i)$	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
ω	Angular frequency in radian/sec
ω_c	Cutoff angular frequency in radian/sec
ω_0	Fundamental angular frequency
ω_s	Angular sampling frequency
w_k	Wiener filter coefficients
$\boldsymbol{w}(m)$	Wiener filter coefficients vector
$W(f)$	Wiener filter frequency response
z	z-transform variable

Abbreviations

ADC	Analogue to digital converter
AR	Autoregressive process
ARMA	Autoregressive moving average process
ATH	Absolute threshold of hearing
ATRAC	Adaptive Transform Acoustic Coding
AWGN	Additive white Gaussian noise
bps	Bits per second
BSS	Blind signal separation
CD	Compact disc
cdf	Cumulative density function
CELP	Code Excited Linear Prediction
Companding	**Comp**ressing Exp**anding**
DAC	Digital to analogue converter
dB	Decibels: 10log10(power ratio) or 10log10(amplitude ratio)
DCT	Discrete cosine transform
Det()	determinant
DFT	Discrete Fourier transform
DNA	Deoxyribonucleic acid
DoA	Direction of arrival
DSP	Digital signal processing
DTW	Dynamic time warping
EM	Estimate-maximise
EM	Electro-magnetic
FFT	Fast Fourier transform
FIR	Finite impulse response
GMM	Gaussian mixture model
GSM	Global system for mobile
HMM	Hidden Markov model
HNM	Harmonic plus noise model
Hz	Unit of frequency in cycles per second
ICA	Independent component analysis
IDCT	Inverse discrete cosine transform
IDFT	Inverse discrete Fourier transform

IFFT	Inverse fast Fourier transform
IID	Independent identically distributed
IIR	Infinite impulse response
ISD	Itakura–Saito distance
ISI	Inter symbol interference
ITU	International Telecommunication Union
JND	Just noticeable distortion
KLT	Karhunen–Loève transform
LF	Liljencrants–Fant
LMS	Least mean squared error
LP	Linear prediction model or Lowpass filter
LPC	Linear prediction coding
LPSS	Spectral subtraction based on linear prediction model
LS	Least square
LSAR	Least square AR interpolation
LSE	Least square error
LSF	Line spectral frequency
LSP	Line spectral pair
LTI	Linear time invariant
MA	Moving average process
MAP	Maximum a posterior estimate
M-ary	Multi-level signalling
MAVE	Minimum absolute value of error estimate
MFCC	Mel frequency cepstral coefficients
MIMO	Multiple input multiple output
ML	Maximum likelihood estimate
MMSE	Minimum mean squared error estimate
MOS	Mean Opinion Score
MP3	MPEG-1 Audio Layer 3
MPEG	Moving Picture Experts Group
ms	Milliseconds
NLMS	Normalised least mean squared error
PCA	Principal component analysis
pdf	Probability density function
PLC	Packet loss concealment
pmf	Probability mass function
PRNG	Pseudo random number generators
psd	Power spectral density
QMF	Quadrature mirror filter
QR	Q is an orthogonal matrix and R is an upper triangular matrix
QRD	Orthogonal matrix decomposition
RF	Radio frequency
RLS	Recursive least square
ROC	Region of convergence
SH or S/H	Sample and hold
SINR	Signal to impulsive noise ratio
SNR	Signal to noise ratio
SPL	Sound pressure level
SQNR	Signal to quantisation noise ratio

std	Standard deviation
STFT	Short time Fourier transform
SVD	Singular value decomposition
ToA	Time of arrival
VAD	Voice activity detector
Var	Variance
ZI	Zero insertion

Contents