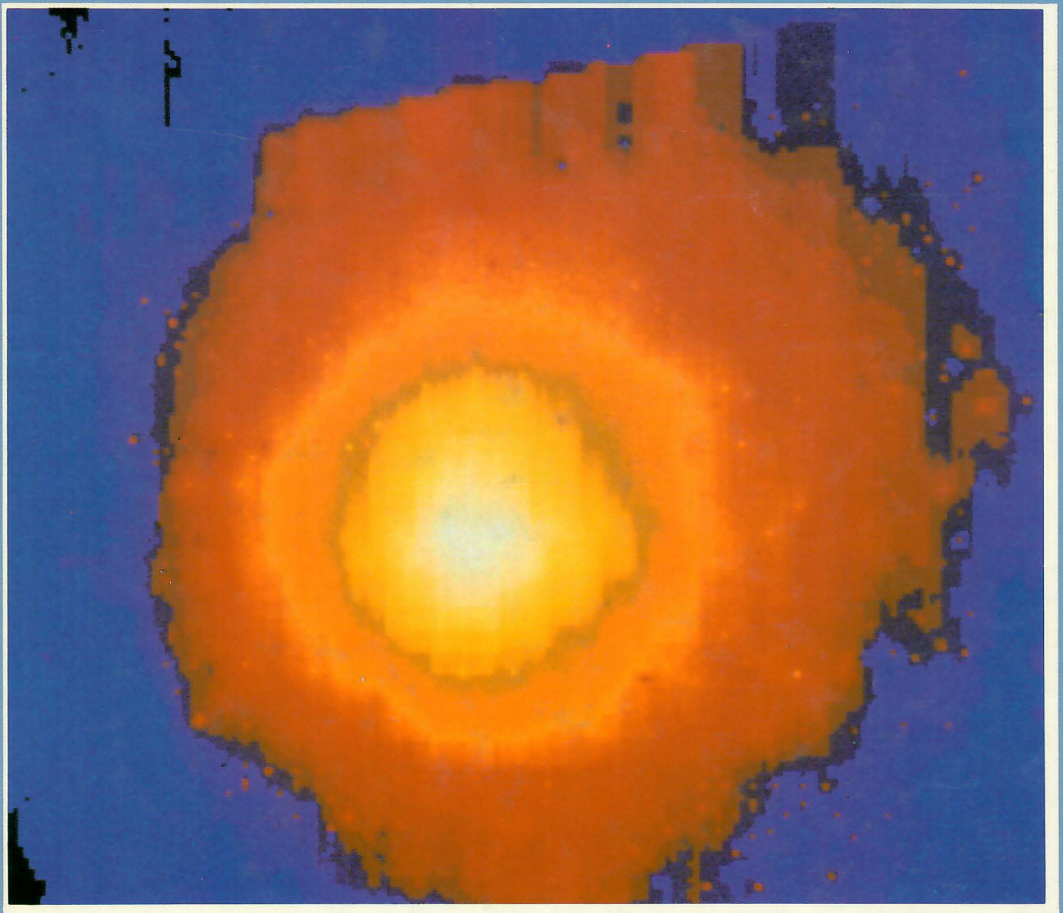

FUNDAMENTALS OF DIGITAL IMAGE PROCESSING—



ANIL K. JAIN

PRENTICE HALL INFORMATION AND SYSTEM SCIENCES SERIES
Thomas Kailath, Series Editor



**PRENTICE HALL INFORMATION
AND SYSTEM SCIENCES SERIES**
Thomas Kailath, Editor

- | | |
|--------------------------------------|---|
| ANDERSON & MOORE | <i>Optimal Control: Linear Quadratic Methods</i> |
| ANDERSON & MOORE | <i>Optimal Filtering</i> |
| ASTROM & WITTENMARK | <i>Computer-Controlled Systems: Theory and Design, 2/E</i> |
| DICKINSON | <i>Systems: Analysis, Design and Computation</i> |
| GARDNER | <i>Statistical Spectral Analysis: A Nonprobabilistic Theory</i> |
| GOODWIN & SIN | <i>Adaptive Filtering, Prediction, and Control</i> |
| GRAY & DAVISSON | <i>Random Processes: A Mathematical Approach for Engineers</i> |
| HAYKIN | <i>Adaptive Filter Theory, 2/E</i> |
| JAIN | <i>Fundamentals of Digital Image Processing</i> |
| JOHNSON | <i>Lectures on Adaptive Parameter Estimation</i> |
| KAILATH | <i>Linear Systems</i> |
| KUMAR & VARAIYA | <i>Stochastic Systems: Estimation, Identification, and Adaptive Control</i> |
| KUNG | <i>VLSI Array Processors</i> |
| KUNG, WHITEHOUSE, &
KAILATH, EDS. | <i>VLSI and Modern Signal Processing</i> |
| KWAKERNAAK & SIVAN | <i>Modern Signals and Systems</i> |
| LANDAU | <i>System Identification and Control Design Using PIM+ Software</i> |
| LIUNG | <i>System Identification: Theory for the User</i> |
| MACOVSKI | <i>Medical Imaging Systems</i> |
| MIDDLETON & GOODWIN | <i>Digital Control and Estimation: A Unified Approach</i> |
| NARENDRA &
ANNASWAMY | <i>Stable Adaptive Systems</i> |
| SASTRY & BODSON | <i>Adaptive Control: Stability, Convergence, and Robustness</i> |
| SOLIMAN & SRINATH | <i>Continuous and Discrete Signals and Systems</i> |
| SPILKER | <i>Digital Communications by Satellite</i> |
| WILLIAMS | <i>Designing Digital Filters</i> |

Library of Congress Cataloging-in-Publication Data

JAIN, ANIL K.

Fundamentals of digital image processing.

Bibliography: p.

Includes index.

1. Image processing—Digital techniques. I. Title.
TA1632.J35 1989 621.36'7 88-12624
ISBN 0-13-336165-9

Editorial/production supervision: Colleen Brosnan

Manufacturing buyer: Mary Noonan

Page layout: Martin Behan

Cover design: Diane Saxe

Logo design: A.M. Bruckstein

Cover art: Halley's comet image by the author

reconstructed from data gathered by NASA's

Pioneer Venus Orbiter in 1986.



© 1989 by Prentice-Hall, Inc.

A Division of Simon & Schuster

Englewood Cliffs, New Jersey 07632

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

Printed in the United States of America

10 9 8 7 6 5

ISBN 0-13-336165-9

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*

SIMON & SCHUSTER ASIA PTE. LTD., *Singapore*

EDITORIA PRENTICE-HALL DO BRASIL, LTDA., *Rio de Janeiro*

Contents

PREFACE	<i>xix</i>
ACKNOWLEDGMENTS	<i>xxi</i>
1 INTRODUCTION	1
1.1 Digital Image Processing: Problems and Applications	1
1.2 Image Representation and Modeling	4
1.3 Image Enhancement	6
1.4 Image Restoration	7
1.5 Image Analysis	7
1.6 Image Reconstruction from Projections	8
1.7 Image Data Compression	9
Bibliography	10
2 TWO-DIMENSIONAL SYSTEMS AND MATHEMATICAL PRELIMINARIES	11
2.1 Introduction	11
2.2 Notation and Definitions	11
2.3 Linear Systems and Shift Invariance	13
2.4 The Fourier Transform	15
<i>Properties of the Fourier Transform, 16</i>	
<i>Fourier Transform of Sequences (Fourier Series), 18</i>	

2.5	The Z-Transform or Laurent Series	20
	<i>Causality and Stability</i> ,	21
2.6	Optical and Modulation Transfer Functions	21
2.7	Matrix Theory Results	22
	<i>Vectors and Matrices</i> ,	22
	<i>Row and Column Ordering</i> ,	23
	<i>Transposition and Conjugation Rules</i> ,	25
	<i>Toeplitz and Circulant Matrices</i> ,	26
	<i>Orthogonal and Unitary Matrices</i> ,	26
	<i>Positive Definiteness and Quadratic Forms</i> ,	27
	<i>Diagonal Forms</i> ,	27
2.8	Block Matrices and Kronecker Products	28
	<i>Block Matrices</i> ,	28
	<i>Kronecker Products</i> ,	30
	<i>Separable Operations</i> ,	31
2.9	Random Signals	31
	<i>Definitions</i> ,	31
	<i>Gaussian or Normal Distribution</i> ,	32
	<i>Gaussian Random Processes</i> ,	32
	<i>Stationary Processes</i> ,	32
	<i>Markov Processes</i> ,	33
	<i>Orthogonality and Independence</i> ,	34
	<i>The Karhunen Loève (KL) Transform</i> ,	34
2.10	Discrete Random Fields	35
	<i>Definitions</i> ,	35
	<i>Separable and Isotropic Covariance Functions</i> ,	36
2.11	The Spectral Density Function	37
	<i>Properties of the SDF</i> ,	38
2.12	Some Results from Estimation Theory	39
	<i>Mean Square Estimates</i> ,	40
	<i>The Orthogonality Principle</i> ,	40
2.13	Some Results from Information Theory	41
	<i>Information</i> ,	42
	<i>Entropy</i> ,	42
	<i>The Rate Distortion Function</i> ,	43
	Problems	44
	Bibliography	47
3	IMAGE PERCEPTION	49
3.1	Introduction	49
3.2	Light, Luminance, Brightness, and Contrast	49
	<i>Simultaneous Contrast</i> ,	51
	<i>Mach Bands</i> ,	53

3.3	MTF of the Visual System	54
3.4	The Visibility Function	55
3.5	Monochrome Vision Models	56
3.6	Image Fidelity Criteria	57
3.7	Color Representation	60
3.8	Color Matching and Reproduction	62
	<i>Laws of Color Matching</i> ,	63
	<i>Chromaticity Diagram</i> ,	65
3.9	Color Coordinate Systems	66
3.10	Color Difference Measures	71
3.11	Color Vision Model	73
3.12	Temporal Properties of Vision	75
	<i>Bloch's Law</i> ,	75
	<i>Critical Fusion Frequency (CFF)</i> ,	75
	<i>Spatial versus Temporal Effects</i> ,	75
	Problems	76
	Bibliography	78

4 IMAGE SAMPLING AND QUANTIZATION

80

4.1	Introduction	80
	<i>Image Scanning</i> ,	80
	<i>Television Standards</i> ,	81
	<i>Image Display and Recording</i> ,	83
4.2	Two-Dimensional Sampling Theory	84
	<i>Bandlimited Images</i> ,	84
	<i>Sampling Versus Replication</i> ,	85
	<i>Reconstruction of the Image from Its</i> <i>Samples</i> ,	85
	<i>Nyquist Rate, Aliasing, and Foldover</i> <i>Frequencies</i> ,	87
	<i>Sampling Theorem</i> ,	88
	<i>Remarks</i> ,	89
4.3	Extensions of Sampling Theory	89
	<i>Sampling Random Fields</i> ,	90
	<i>Sampling Theorem for Random Fields</i> ,	90
	<i>Remarks</i> ,	90
	<i>Nonrectangular Grid Sampling and</i> <i>Interlacing</i> ,	91
	<i>Hexagonal Sampling</i> ,	92
	<i>Optimal Sampling</i> ,	92

4.4	Practical Limitations in Sampling and Reconstruction 93
	<i>Sampling Aperture, 93</i>
	<i>Display Aperture/Interpolation Function, 94</i>
	<i>Lagrange Interpolation, 98</i>
	<i>Moire Effect and Flat Field Response, 99</i>
4.5	Image Quantization 99
4.6	The Optimum Mean Square or Lloyd-Max Quantizer 101
	<i>The Uniform Optimal Quantizer, 103</i>
	<i>Properties of the Optimum Mean Square Quantizer, 103</i>
	<i>Proofs, 112</i>
4.7	A Compandor Design 113
	<i>Remarks, 114</i>
4.8	The Optimum Mean Square Uniform Quantizer for Nonuniform Densities 115
4.9	Examples, Comparison, and Practical Limitations 115
4.10	Analytic Models for Practical Quantizers 118
4.11	Quantization of Complex Gaussian Random Variables 119
4.12	Visual Quantization 119
	<i>Contrast Quantization, 120</i>
	<i>Pseudorandom Noise Quantization, 120</i>
	<i>Halftone Image Generation, 121</i>
	<i>Color Quantization, 122</i>
	Problems 124
	Bibliography 128

5 IMAGE TRANSFORMS

132

5.1	Introduction 132
5.2	Two-Dimensional Orthogonal and Unitary Transforms 134
	<i>Separable Unitary Transforms, 134</i>
	<i>Basis Images, 135</i>
	<i>Kronecker Products and Dimensionality, 137</i>
	<i>Dimensionality of Image Transforms, 138</i>
	<i>Transform Frequency, 138</i>
	<i>Optimum Transform, 138</i>
5.3	Properties of Unitary Transforms 138
	<i>Energy Conservation and Rotation, 138</i>
	<i>Energy Compaction and Variances of Transform Coefficients, 139</i>
	<i>Decorrelation, 140</i>
	<i>Other Properties, 140</i>

5.4	The One-Dimensional Discrete Fourier Transform (DFT)	141
	<i>Properties of the DFT/Unitary DFT, 141</i>	
5.5	The Two-Dimensional DFT	145
	<i>Properties of the Two-Dimensional DFT, 147</i>	
5.6	The Cosine Transform	150
	<i>Properties of the Cosine Transform, 151</i>	
5.7	The Sine Transform	154
	<i>Properties of the Sine Transform, 154</i>	
5.8	The Hadamard Transform	155
	<i>Properties of the Hadamard Transform, 157</i>	
5.9	The Haar Transform	159
	<i>Properties of the Haar Transform, 161</i>	
5.10	The Slant Transform	161
	<i>Properties of the Slant Transform, 162</i>	
5.11	The KL Transform	163
	<i>KL Transform of Images, 164</i>	
	<i>Properties of the KL Transform, 165</i>	
5.12	A Sinusoidal Family of Unitary Transforms	175
	<i>Approximation to the KL Transform, 176</i>	
5.13	Outer Product Expansion and Singular Value Decomposition	176
	<i>Properties of the SVD Transform, 177</i>	
5.14	Summary	180
	Problems	180
	Bibliography	185

6 IMAGE REPRESENTATION BY STOCHASTIC MODELS

189

6.1	Introduction	189
	<i>Covariance Models, 189</i>	
	<i>Linear System Models, 189</i>	
6.2	One-Dimensional Causal Models	190
	<i>Autoregressive (AR) Models, 190</i>	
	<i>Properties of AR Models, 191</i>	
	<i>Application of AR Models in Image Processing,</i>	
	<i>193</i>	
	<i>Moving Average (MA) Representations, 194</i>	
	<i>Autoregressive Moving Average (ARMA)</i>	
	<i>Representations, 195</i>	
	<i>State Variable Models, 195</i>	
	<i>Image Scanning Models, 196</i>	

6.3	One-Dimensional Spectral Factorization	196
	<i>Rational SDFs</i> ,	197
	<i>Remarks</i> ,	198
6.4	AR Models, Spectral Factorization, and Levinson Algorithm	198
	<i>The Levinson-Durbin Algorithm</i> ,	198
6.5	Noncausal Representations	200
	<i>Remarks</i> ,	201
	<i>Noncausal MVRs for Autoregressive Sequences</i> ,	201
	<i>A Fast KL Transform</i> ,	202
	<i>Optimum Interpolation of Images</i> ,	204
6.6	Linear Prediction in Two Dimensions	204
	<i>Causal Prediction</i> ,	205
	<i>Semicausal Prediction</i> ,	206
	<i>Noncausal Prediction</i> ,	206
	<i>Minimum Variance Prediction</i> ,	206
	<i>Stochastic Representation of Random Fields</i> ,	207
	<i>Finite-Order MVRs</i> ,	208
	<i>Remarks</i> ,	209
	<i>Stability of Two-Dimensional Systems</i> ,	212
6.7	Two-Dimensional Spectral Factorization and Spectral Estimation Via Prediction Models	213
	<i>Separable Models</i> ,	214
	<i>Realization of Noncausal MVRs</i> ,	215
	<i>Realization of Causal and Semicausal MVRs</i> ,	216
	<i>Realization via Orthogonality Condition</i> ,	216
6.8	Spectral Factorization via the Wiener-Doob Homomorphic Transformation	219
	<i>Causal MVRs</i> ,	220
	<i>Semicausal WNDRs</i> ,	220
	<i>Semicausal MVRs</i> ,	222
	<i>Remarks and Examples</i> ,	222
6.9	Image Decomposition, Fast KL Transforms, and Stochastic Decoupling	223
	<i>Periodic Random Fields</i> ,	223
	<i>Noncausal Models and Fast KL Transforms</i> ,	224
	<i>Semicausal Models and Stochastic Decoupling</i> ,	225
6.10	Summary	226
	Problems	227
	Bibliography	230

- 7.1 Introduction 233
- 7.2 Point Operations 235
Contrast Stretching, 235
Clipping and Thresholding, 235
Digital Negative, 238
Intensity Level Slicing, 238
Bit Extraction, 239
Range Compression, 240
Image Subtraction and Change Detection, 240
- 7.3 Histogram Modeling 241
Histogram Equalization, 241
Histogram Modification, 242
Histogram Specification, 243
- 7.4 Spatial Operations 244
Spatial Averaging and Spatial Low-pass Filtering, 244
Directional Smoothing, 245
Median Filtering, 246
Other Smoothing Techniques, 249
Unsharp Masking and Crispensing, 249
Spatial Low-pass, High-pass and Band-pass Filtering, 250
Inverse Contrast Ratio Mapping and Statistical Scaling, 252
Magnification and Interpolation (Zooming), 253
Replication, 253
Linear Interpolation, 253
- 7.5 Transform Operations 256
Generalized Linear Filtering, 256
Root Filtering, 258
Generalized Cepstrum and Homomorphic Filtering, 259
- 7.6 Multispectral Image Enhancement 260
Intensity Ratios, 260
Log-Ratios, 261
Principal Components, 261
- 7.7 False Color and Pseudocolor 262
- 7.8 Color Image Enhancement 262
- 7.9 Summary 263
Problems 263
Bibliography 265

- 8.1 Introduction 267
- 8.2 Image Observation Models 268
 - Image Formation Models*, 269
 - Detector and Recorder Models*, 273
 - Noise Models*, 273
 - Sampled Image Observation Models*, 275
- 8.3 Inverse and Wiener Filtering 275
 - Inverse Filter*, 275
 - Pseudoinverse Filter*, 276
 - The Wiener Filter*, 276
 - Remarks*, 279
- 8.4 Finite Impulse Response (FIR) Wiener Filters 284
 - Filter Design*, 284
 - Remarks*, 285
 - Spatially Varying FIR Filters*, 287
- 8.5 Other Fourier Domain Filters 290
 - Geometric Mean Filter*, 291
 - Nonlinear Filters*, 291
- 8.6 Filtering Using Image Transforms 292
 - Wiener Filtering*, 292
 - Remarks*, 293
 - Generalized Wiener Filtering*, 293
 - Filtering by Fast Decompositions*, 295
- 8.7 Smoothing Splines and Interpolation 295
 - Remarks*, 297
- 8.8 Least Squares Filters 297
 - Constrained Least Squares Restoration*, 297
 - Remarks*, 298
- 8.9 Generalized Inverse, SVD, and Iterative Methods 299
 - The Pseudoinverse*, 299
 - Minimum Norm Least Squares (MNLS)*
 - Solution and the Generalized Inverse*, 300
 - One-step Gradient Methods*, 301
 - Van Cittert Filter*, 301
 - The Conjugate Gradient Method*, 302
 - Separable Point Spread Functions*, 303
- 8.10 Recursive Filtering For State Variable Systems 304
 - Kalman Filtering*, 304
 - Remarks*, 307

8.11	Causal Models and Recursive Filtering	307
	<i>A Vector Recursive Filter, 308</i>	
	<i>Stationary Models, 310</i>	
	<i>Steady-State Filter, 310</i>	
	<i>A Two-Stage Recursive Filter, 310</i>	
	<i>A Reduced Update Filter, 310</i>	
	<i>Remarks, 311</i>	
8.12	Semicausal Models and Semirecursive Filtering	311
	<i>Filter Formulation, 312</i>	
8.13	Digital Processing of Speckle Images	313
	<i>Speckle Representation, 313</i>	
	<i>Speckle Reduction: N-Look Method, 315</i>	
	<i>Spatial Averaging of Speckle, 315</i>	
	<i>Homomorphic Filtering, 315</i>	
8.14	Maximum Entropy Restoration	316
	<i>Distribution-Entropy Restoration, 317</i>	
	<i>Log-Entropy Restoration, 318</i>	
8.15	Bayesian Methods	319
	<i>Remarks, 320</i>	
8.16	Coordinate Transformation and Geometric Correction	320
8.17	Blind Deconvolution	322
8.18	Extrapolation of Bandlimited Signals	323
	<i>Analytic Continuation, 323</i>	
	<i>Super-resolution, 323</i>	
	<i>Extrapolation Via Prolate Spheroidal Wave Functions (PSWFs), 324</i>	
	<i>Extrapolation by Error Energy Reduction, 324</i>	
	<i>Extrapolation of Sampled Signals, 326</i>	
	<i>Minimum Norm Least Squares (MNLS)</i>	
	<i>Extrapolation, 326</i>	
	<i>Iterative Algorithms, 327</i>	
	<i>Discrete Prolate Spheroidal Sequences (DPSS), 327</i>	
	<i>Mean Square Extrapolation, 328</i>	
	<i>Generalization to Two Dimensions, 328</i>	
8.19	Summary	330
	Problems	331
	Bibliography	335
9	IMAGE ANALYSIS AND COMPUTER VISION	342
9.1	Introduction	342
9.2	Spatial Feature Extraction	344
	<i>Amplitude Features, 344</i>	
	<i>Histogram Features, 344</i>	

9.3	Transform Features	346
9.4	Edge Detection	347
	<i>Gradient Operators</i> ,	348
	<i>Compass Operators</i> ,	350
	<i>Laplace Operators and Zero Crossings</i> ,	351
	<i>Stochastic Gradients</i> ,	353
	<i>Performance of Edge Detection Operators</i> ,	355
	<i>Line and Spot Detection</i> ,	356
9.5	Boundary Extraction	357
	<i>Connectivity</i> ,	357
	<i>Contour Following</i> ,	358
	<i>Edge Linking and Heuristic Graph Searching</i> ,	358
	<i>Dynamic Programming</i> ,	359
	<i>Hough Transform</i> ,	362
9.6	Boundary Representation	362
	<i>Chain Codes</i> ,	363
	<i>Fitting Line Segments</i> ,	364
	<i>B-Spline Representation</i> ,	364
	<i>Fourier Descriptors</i> ,	370
	<i>Autoregressive Models</i> ,	374
9.7	Region Representation	375
	<i>Run-length Codes</i> ,	375
	<i>Quad-Trees</i> ,	375
	<i>Projections</i> ,	376
9.8	Moment Representation	377
	<i>Definitions</i> ,	377
	<i>Moment Representation Theorem</i> ,	378
	<i>Moment Matching</i> ,	378
	<i>Orthogonal Moments</i> ,	379
	<i>Moment Invariants</i> ,	380
	<i>Applications of Moment Invariants</i> ,	381
9.9	Structure	381
	<i>Medial Axis Transform</i> ,	381
	<i>Morphological Processing</i> ,	384
	<i>Morphological Transforms</i> ,	387
9.10	Shape Features	390
	<i>Geometry Features</i> ,	391
	<i>Moment-Based Features</i> ,	392
9.11	Texture	394
	<i>Statistical Approaches</i> ,	394
	<i>Structural Approaches</i> ,	398
	<i>Other Approaches</i> ,	399

9.12	Scene Matching and Detection 400
	<i>Image Subtraction, 400</i>
	<i>Template Matching and Area Correlation, 400</i>
	<i>Matched Filtering, 403</i>
	<i>Direct Search Methods, 404</i>
9.13	Image Segmentation 407
	<i>Amplitude Thresholding or Window Slicing, 407</i>
	<i>Component Labeling, 409</i>
	<i>Boundary-based Approaches, 411</i>
	<i>Region-based Approaches and Clustering, 412</i>
	<i>Template Matching, 413</i>
	<i>Texture Segmentation, 413</i>
9.14	Classification Techniques 414
	<i>Supervised Learning, 414</i>
	<i>Nonsupervised Learning or Clustering, 418</i>
9.15	Image Understanding 421
	Problems 422
	Bibliography 425

10 IMAGE RECONSTRUCTION FROM PROJECTIONS

431

10.1	Introduction 431
	<i>Transmission Tomography, 431</i>
	<i>Reflection Tomography, 432</i>
	<i>Emission Tomography, 433</i>
	<i>Magnetic Resonance Imaging, 434</i>
	<i>Projection-based Image Processing, 434</i>
10.2	The Radon Transform 434
	<i>Definition, 434</i>
	<i>Notation, 436</i>
	<i>Properties of the Radon Transform, 437</i>
10.3	The Back-projection Operator 439
	<i>Definition, 439</i>
	<i>Remarks, 440</i>
10.4	The Projection Theorem 442
	<i>Remarks, 443</i>
10.5	The Inverse Radon Transform 444
	<i>Remarks, 445</i>
	<i>Convolution Back-projection Method, 446</i>
	<i>Filter Back-projection Method, 446</i>
	<i>Two-Dimensional Filtering via the Radon Transform, 447</i>

10.6	Convolution/Filter Back-projection Algorithms: Digital Implementation 448
	<i>Sampling Considerations, 448</i>
	<i>Choice of Filters, 448</i>
	<i>Convolution Back-projection Algorithm, 450</i>
	<i>Filter Back-projection Algorithm, 451</i>
	<i>Reconstruction Using a Parallel Pipeline Processor, 452</i>
10.7	Radon Transform of Random Fields 452
	<i>A Unitary Transform $\hat{R}$, 452</i>
	<i>Radon Transform Properties for Random Fields, 456</i>
	<i>Projection Theorem for Random Fields, 457</i>
10.8	Reconstruction from Blurred Noisy Projections 458
	<i>Measurement Model, 458</i>
	<i>The Optimum Mean Square Filter, 458</i>
	<i>Remarks, 458</i>
10.9	Fourier Reconstruction 462
	<i>Algorithm, 462</i>
	<i>Reconstruction of Magnetic Resonance Images, 463</i>
10.10	Fan-Beam Reconstruction 464
10.11	Algebraic Methods 465
	<i>The Reconstruction Problem as a Set of Linear Equations, 465</i>
	<i>Algebraic Reconstruction Techniques, 466</i>
10.12	Three-Dimensional Tomography 468
	<i>Three-Dimensional Reconstruction Algorithms, 469</i>
10.13	Summary 470
	Problems 470
	Bibliography 473

11 IMAGE DATA COMPRESSION

476

11.1	Introduction 476
	<i>Image Raw Data Rates, 476</i>
	<i>Data Compression versus Bandwidth Compression, 477</i>
	<i>Information Rates, 477</i>
	<i>Subsampling, Coarse Quantization, Frame Repetition, and Interlacing, 479</i>

11.2	Pixel Coding 479
	PCM, 480
	Entropy Coding, 480
	Run-Length Coding, 481
	Bit-Plane Encoding, 483
11.3	Predictive Techniques 483
	Basic Principle, 483
	Feedback versus Feedforward Prediction, 484
	Distortionless Predictive Coding, 485
	Performance Analysis of DPCM, 486
	Delta Modulation, 488
	Line-by-Line DPCM, 490
	Two-Dimensional DPCM, 491
	Performance Comparisons, 493
	Remarks, 494
	Adaptive Techniques, 495
	Other Methods, 497
11.4	Transform Coding Theory 498
	The Optimum Transform Coder, 498
	Proofs, 499
	Remarks, 501
	Bit Allocation and Rate-Distortion
	Characteristics, 501
11.5	Transform Coding of Images 504
	Two-Dimensional Coding Algorithm, 504
	Transform Coding Performances Trade-offs and
	Examples, 507
	Zonal versus Threshold Coding, 508
	Fast KL Transform Coding, 510
	Remarks, 512
	Two-Source Coding, 513
	Transform Coding under Visual Criterion, 515
	Adaptive Transform Coding, 515
	Summary of Transform Coding, 516
11.6	Hybrid Coding and Vector DPCM 518
	Basic Idea, 518
	Adaptive Hybrid Coding, 520
	Hybrid Coding Conclusions, 521
11.7	Interframe Coding 521
	Frame Repetition, 521
	Resolution Exchange, 521
	Conditional Replenishment, 522
	Adaptive Predictive Coding, 522
	Predictive Coding with Motion Compensation,
	524
	Interframe Hybrid Coding, 527
	Three-Dimensional Transform Coding, 529

11.8	Image Coding in the Presence of Channel Errors	532
	<i>The Optimum Mean Square Decoder, 532</i>	
	<i>The Optimum Encoding Rule, 533</i>	
	<i>Optimization of PCM Transmission, 534</i>	
	<i>Channel Error Effects in DPCM, 536</i>	
	<i>Optimization of Transform Coding, 537</i>	
11.9	Coding of Two Tone Images	540
	<i>Run-length Coding, 540</i>	
	<i>White Block Skipping, 546</i>	
	<i>Prediction Differential Quantization, 547</i>	
	<i>Relative Address Coding, 547</i>	
	<i>CCITT Modified Relative Element Address</i>	
	<i>Designate Coding, 548</i>	
	<i>Predictive Coding, 551</i>	
	<i>Adaptive Predictors, 552</i>	
	<i>Comparison of Algorithms, 553</i>	
	<i>Other Methods, 553</i>	
11.10	Color and Multispectral Image Coding	553
11.11	Summary	557
	Problems	557
	Bibliography	561

INDEX

566

This is often used as the covariance model of a scan line of monochrome images. For an $N \times 1$ vector $\mathbf{u} = \{u(n), 1 \leq n \leq N\}$, its covariance matrix is $\{r(m-n)\}$, i.e.,

$$\mathbf{R} = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{N-1} \\ \rho & 1 & \rho & \dots & \rho^{N-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho^{N-1} & \dots & \rho & 1 & \end{bmatrix} \quad (2.68)$$

which is Toeplitz. In fact the covariance and autocorrelation matrices of any stationary sequence are Toeplitz. Conversely, any sequence, finite or infinite, can be called stationary if its covariance and autocorrelation matrices are Toeplitz.

Orthogonality and Independence

Two random variables x and y are called *independent* if and only if their joint probability density function is a product of their marginal densities, i.e.,

$$p_{x,y}(x,y) = p_x(x)p_y(y) \quad (2.69)$$

Two random sequences $x(n)$ and $y(n)$ are called *independent* if and only if for every n and n' , the random variables $x(n)$ and $y(n')$ are independent.

The random variables x and y are said to be *orthogonal* if

$$E[xy^*] = 0 \quad (2.70)$$

and are called *uncorrelated* if

$$E[xy^*] = (E[x])(E[y^*])$$

or

$$E[(x - \mu_x)(y - \mu_y)^*] = 0 \quad (2.71)$$

Thus zero mean uncorrelated random variables are also orthogonal. Gaussian random variables which are uncorrelated are also independent.

The Karhunen-Loève (KL) Transform

Let $\{x(n), 1 \leq n \leq N\}$ be a complex random sequence whose autocorrelation matrix is $\mathbf{R}$. Let Φ be an $N \times N$ unitary matrix, which reduces $\mathbf{R}$ to its diagonal form Λ [see (2.44)]. The transformed vector

$$\mathbf{y} = \Phi^{*T} \mathbf{x} \quad (2.72)$$

is called the Karhunen-Loève (KL) transform of $\mathbf{x}$. It satisfies the property

$$\begin{aligned} E[\mathbf{y}\mathbf{y}^{*T}] &= \Phi^{*T} \{E[\mathbf{x}\mathbf{x}^{*T}]\} \Phi = \Phi^{*T} \mathbf{R} \Phi = \Lambda \\ \Rightarrow E[y(k)y^*(l)] &= \lambda_k \delta(k-l) \end{aligned} \quad (2.73)$$

i.e., the elements of the transformed sequence $y(k)$ are *orthogonal*. If $\mathbf{R}$ represents the covariance matrix rather than the autocorrelation matrix of $\mathbf{x}$, then the sequence $y(k)$ is *uncorrelated*. The unitary matrix Φ^{*T} is called the KL transform matrix. Its rows are the conjugate eigenvectors of $\mathbf{R}$, i.e., it is the conjugate transpose of the eigenmatrix of $\mathbf{R}$. The KL transform is of fundamental importance in digital signal and image processing. Its applications and properties are considered in Chapter 5.

2.10 DISCRETE RANDOM FIELDS

In statistical representation of images, each pixel is considered as a random variable. Thus we think of a given image as a sample function of an ensemble of images. Such an ensemble would be adequately defined by a joint probability density of the array of random variables. For practical image sizes, the number of random variables is very large (262,144 for 512×512 images). Thus it is difficult to specify a realistic joint density function because it would be an enormous task to measure it. One possibility is to specify the ensemble by its first- and second-order moments only (mean and covariances). Even with this simplifying constraint, the task of determining realistic model parameters remains difficult. Various approaches for stochastic modeling are considered in Chapter 6. Here we consider some basic definitions that will be useful in the subsequent chapters.

Definitions

When each sample of a two-dimensional sequence is a random variable, we call it a *discrete random field*. When the random field represents an ensemble of images (such as television images or satellite images), we call it a *random image*. The term *random field* will apply to any two-dimensional random sequence.

The mean and covariance functions of a complex random field are defined as

$$E[u(m, n)] = \mu(m, n) \quad (2.74)$$

$$\begin{aligned} \text{Cov}[u(m, n), u(m', n')] &\triangleq E[(u(m, n) - \mu(m, n))(u^*(m', n') - \mu^*(m', n'))] \\ &= r_u(m, n; m', n') = r(m, n; m', n') \end{aligned} \quad (2.75)$$

Often, we will consider the stationary case where

$$\begin{aligned} \mu(m, n) &= \mu = \text{constant} \\ r_u(m, n; m', n') &= r_u(m - m', n - n') = r(m - m', n - n') \end{aligned} \quad (2.76)$$

As before, whenever there is no confusion, we will drop the subscript u from r_u . A random field satisfying (2.76) is also called *shift invariant*, *translational* (or *spatial*) *invariant*, *homogeneous*, or *wide-sense stationary*. Unless otherwise mentioned, stationarity will always be implied in the wide sense.

We will denote the covariance function of a stationary random field $u(m, n)$ by $r_u(m, n)$ or $r(m, n)$, implying that

$$r(m, n) = \text{Cov}[u(m, n), u(0, 0)] = \text{Cov}[u(m' + m, n' + n), u(m', n')] \quad \forall(m', n') \quad (2.77)$$

Nyquist sampling intervals would be $\Delta x = \Delta y \triangleq \Delta_1 = 1$. If the sampling grid G_2 is chosen, which is a 45° rotation of G_1 but with intersample distance of Δ_2 , the spectrum of the sampled image will repeat on a grid similar to G_2 (with spacing $1/\Delta_2$) (Fig. 4.9e). Therefore, if $\Delta_2 = \sqrt{2}$, there will be no aliasing, but the sampling density has been reduced by half. Thus if an image does not contain the high frequencies in both the dimensions simultaneously, then its sampling rate can be reduced by a factor of 2. This theory is used in line interlacing television signals because the human vision is insensitive to high spatial frequencies in areas of large motion (high temporal frequencies). The interlaced television signal can be considered as a three-dimensional signal $f(x, y, t)$ sampled in vertical (y) and temporal (t) dimensions. If ξ_1 and ξ_2 represent the temporal and vertical frequencies, respectively, then Fig. 4.9e represents the projection in (ξ_1, ξ_2) plane of the three-dimensional spectrum of the interlaced television signal.

In digital television, all three coordinates x , y , and t are sampled. The preceding interlacing concept can be extended to yield the *line quincunx* sampling pattern [10]. Here each field uses an interlaced grid as in Fig. 4.9c, which reduces the sampling rate by another factor of two.

Hexagonal Sampling

For functions that are circularly symmetric and/or bandlimited over a circular region, it can be shown that sampling on a hexagonal lattice requires 13.4 percent fewer samples than rectangular sampling. Alternatively, for the same sampling rate less aliasing is obtained on a hexagonal lattice than a rectangular lattice. Details are available in [14].

Optimal Sampling

Equation (4.16) provides the interpretation that the sampling process transforms a *continuous function* $f(x, y)$ into a *sequence* $f(m\Delta x, n\Delta y)$ from which the original function can be recovered. Therefore, the coefficients of any convergent series expansion of $f(x, y)$ can be considered to give a generalized form of sampling. Such sampling is not restricted to bandlimited functions. For bandlimited functions the sinc functions are optimal for recovering the original function $f(x, y)$ from the samples $f(m\Delta x, n\Delta y)$. For bandlimited random fields, the reconstructed random field converges to the original in the mean square sense.

More generally, there are functions that are optimal in the sense that they sample a random image to give a finite sequence such that the mean square error between the original and the reconstructed images is minimized. In particular, a series expansion of special interest is

$$f(x, y) = \sum_{m,n=0}^{\infty} a_{m,n} \phi_{m,n}(x, y) \quad (4.24)$$

where $\{\phi_{m,n}(x, y)\}$ are the eigenfunctions of the autocorrelation function of the random field $f(x, y)$. This is called the *Karhunen-Loève (KL) series expansion* of the random field. This expansion is such that $a_{m,n}$ are orthogonal random variables,

and, for a given number of terms, the mean square error in the reconstructed image is minimum among all possible sampling functions. This property is useful in developing data compression techniques for images.

The main difficulty in utilizing the preceding result for optimal sampling of practical (finite size) images is in generating the coefficients $a_{m,n}$. In conventional sampling (via the sinc functions), the coefficients $a_{m,n}$ are simply the values $f(m\Delta x, n\Delta y)$, which are easy to obtain. Nevertheless, the theory of KL expansion is useful in determining bounds on performance and serves as an important guide in the design of many image processing algorithms.

4.4 PRACTICAL LIMITATIONS IN SAMPLING AND RECONSTRUCTION

The foregoing sampling theory is based on several idealizations. Real-world images are not bandlimited, which means aliasing errors occur. These can be reduced by low-pass filtering the input image prior to sampling but at the cost of attenuating higher spatial frequencies. Such resolution loss, which results in blurring of the image, also occurs because practical scanners have finite apertures. Finally, the reconstruction system can never be the ideal low-pass filter required by the sampling theory. Its transfer function depends on the display aperture. Figure 4.10 represents the practical sampling/reconstruction systems.

Sampling Aperture

A practical sampling system gives an output $g_s(x, y)$, which can be modeled as (see Fig. 4.10)

$$\begin{aligned} g(x, y) &\triangleq p_s(x, y) \star f(x, y) = p_s(-x, -y) \otimes f(x, y) \\ &= \iint_{\mathcal{A}} p_s(x' - x, y' - y) f(x', y') dx' dy' \end{aligned} \quad (4.25)$$

$$g_s(x, y) = \text{comb}(x, y; \Delta x, \Delta y) g(x, y) \quad (4.26)$$

where $p_s(x, y)$ denotes the light distribution in the aperture and $\mathcal{A}$ denotes its shape. In practice the aperture is symmetric with respect to 180° rotation, that is, $p_s(x, y) = p_s(-x, -y)$. Equation (4.25) is the spatial correlation of $f(x, y)$ with

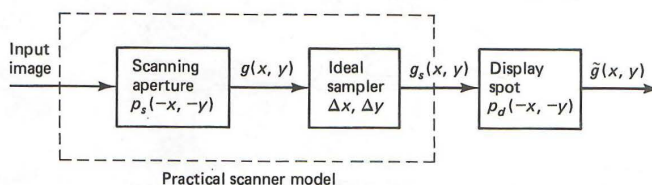


Figure 4.10 Practical sampling and reconstruction. In the ideal case $p_s(x, y) = p_d(x, y) = \delta(x, y)$.

$$i = \frac{N}{2} + 1, \quad \text{sequency} = 3$$

$$\frac{N}{2} + 2 \leq i \leq N - 1, \quad \text{sequency} = \begin{cases} 2\left(i - \frac{N}{2}\right) + 1, & i = \text{even} \\ 2\left(i - \frac{N}{2}\right), & i = \text{odd} \end{cases}$$

5.11 THE KL TRANSFORM

The KL transform was originally introduced as a series expansion for continuous random processes by Karhunen [27] and Loeve [28]. For random sequences Hotelling [26] first studied what was called a method of principal components, which is the discrete equivalent of the KL series expansion. Consequently, the KL transform is also called the Hotelling transform or the method of principal components.

For a real $N \times 1$ random vector $\mathbf{u}$, the basis vectors of the KL transform (see Section 2.9) are given by the orthonormalized eigenvectors of its autocorrelation matrix $\mathbf{R}$, that is,

$$\mathbf{R}\boldsymbol{\phi}_k = \lambda_k \boldsymbol{\phi}_k, \quad 0 \leq k \leq N - 1 \quad (5.132)$$

The KL transform of $\mathbf{u}$ is defined as

$$\mathbf{v} = \boldsymbol{\Phi}^* \mathbf{u} \quad (5.133)$$

and the inverse transform is

$$\mathbf{u} = \boldsymbol{\Phi} \mathbf{v} = \sum_{k=0}^{N-1} v(k) \boldsymbol{\phi}_k \quad (5.134)$$

where $\boldsymbol{\phi}_k$ is the k th column of $\boldsymbol{\Phi}$. From (2.44) we know $\boldsymbol{\Phi}$ reduces $\mathbf{R}$ to its diagonal form, that is,

$$\boldsymbol{\Phi}^* \mathbf{R} \boldsymbol{\Phi} = \boldsymbol{\Lambda} = \text{Diag}\{\lambda_k\} \quad (5.135)$$

We often work with the covariance matrix rather than the autocorrelation matrix. With $\boldsymbol{\mu} \triangleq E[\mathbf{u}]$, then

$$\mathbf{R}_0 \triangleq \text{cov}[\mathbf{u}] \triangleq E[(\mathbf{u} - \boldsymbol{\mu})(\mathbf{u} - \boldsymbol{\mu})^T] = E[\mathbf{u}\mathbf{u}^T] - \boldsymbol{\mu}\boldsymbol{\mu}^T = \mathbf{R} - \boldsymbol{\mu}\boldsymbol{\mu}^T \quad (5.136)$$

If the vector $\boldsymbol{\mu}$ is known, then the eigenmatrix of $\mathbf{R}_0$ determines the KL transform of the zero mean random process $\mathbf{u} - \boldsymbol{\mu}$. In general, the KL transform of $\mathbf{u}$ and $\mathbf{u} - \boldsymbol{\mu}$ need not be identical.

Note that whereas the image transforms considered earlier were functionally independent of the data, the KL transform depends on the (second-order) statistics of the data.

Example 5.4 (KL Transform of Markov-1 Sequences)

The covariance matrix of a zero mean Markov sequence of N elements is given by (2.68). Its eigenvalues λ_k and the eigenvectors $\boldsymbol{\phi}_k$ are given by

$$\lambda_k = \frac{1 - \rho^2}{1 - 2\rho \cos \omega_k + \rho^2}$$

$$\phi_k(m) = \phi(m, k) \quad (5.137)$$

$$= \left(\frac{2}{N + \lambda_k} \right)^{1/2} \sin \left(\omega_k \left(m + 1 - \frac{N+1}{2} \right) + \frac{(k+1)\pi}{2} \right), \quad 0 \leq m, k \leq N-1$$

where the $\{\omega_k\}$ are the positive roots of the equation

$$\tan(N\omega) = -\frac{(1-\rho^2) \sin \omega}{\cos \omega - 2\rho + \rho^2 \cos \omega}, \quad N \text{ even} \quad (5.138)$$

A similar result holds when N is odd. This is a transcendental equation that gives rise to nonharmonic sinusoids $\phi_k(m)$. Figure 5.1 shows the basis vectors of this 8×8 KL transform for $\rho = 0.95$. Note the basis vectors of the KLT and the DCT are quite similar. Because $\phi_k(m)$ are nonharmonic, a fast algorithm for this transform does not exist. Also note, the KL transform matrix is $\Phi^T \triangleq \{\phi(k, m)\}$.

Example 5.5

Since the unitary DFT reduces any circulant matrix to a diagonal form, it is the KL transform of all random sequences with circulant autocorrelation matrices, that is, for all periodic random sequences.

The DCT is the KL transform of a random sequence whose autocorrelation matrix $\mathbf{R}$ commutes with $\mathbf{Q}_c$ of (5.95) (that is, if $\mathbf{R}\mathbf{Q}_c = \mathbf{Q}_c\mathbf{R}$). Similarly, the DST is the KL transform of all random sequences whose autocorrelation matrices commute with $\mathbf{Q}$ of (5.102).

KL Transform of Images

If an $N \times N$ image $u(m, n)$ is represented by a random field whose autocorrelation function is given by

$$E[u(m, n)u(m', n')] = r(m, n; m', n'), \quad 0 \leq m, m', n, n' \leq N-1 \quad (5.139)$$

then the basis images of the KL transform are the orthonormalized eigenfunctions $\psi_{k,l}(m, n)$ obtained by solving

$$\begin{aligned} \sum_{m'=0}^{N-1} \sum_{n'=0}^{N-1} r(m, n; m', n') \psi_{k,l}(m', n') \\ = \lambda_{k,l} \psi_{k,l}(m, n), \quad 0 \leq k, l \leq N-1, 0 \leq m, n \leq N-1 \end{aligned} \quad (5.140)$$

In matrix notation this can be written as

$$\mathcal{R}\psi_i = \lambda_i \psi_i, \quad i = 0, \dots, N^2 - 1 \quad (5.141)$$

where ψ_i is an $N^2 \times 1$ vector representation of $\psi_{k,l}(m, n)$ and $\mathcal{R}$ is an $N^2 \times N^2$ autocorrelation matrix of the image mapped into an $N^2 \times 1$ vector u . Thus

$$\mathcal{R} = E[uu^T] \quad (5.142)$$

If $\mathcal{R}$ is separable, then the $N^2 \times N^2$ matrix Ψ whose columns are $\{\psi_i\}$ becomes separable (see Table 2.7). For example, let

$$r(m, n; m', n') = r_1(m, m')r_2(n, n') \quad (5.143)$$

$$\psi_{k,l}(m, n) = \phi_1(m, k)\phi_2(n, l) \quad (5.144)$$

In matrix notation this means

$$\mathcal{R} = \mathbf{R}_1 \otimes \mathbf{R}_2, \quad \Psi = \Phi_1 \otimes \Phi_2 \quad (5.145)$$

where

$$\Phi_j \mathbf{R}_j \Phi_j^{*T} = \Lambda_j, \quad j = 1, 2 \quad (5.146)$$

and the KL transform of u is

$$v = \Psi^{*T} u = [\Phi_1^{*T} \otimes \Phi_2^{*T}] u \quad (5.147)$$

For row-ordered vectors this is equivalent to

$$\mathbf{V} = \Phi_1^{*T} \mathbf{U} \Phi_2^{*T} \quad (5.148)$$

and the inverse KL transform is

$$\mathbf{U} = \Phi_1 \mathbf{V} \Phi_2^T \quad (5.149)$$

The advantage in modeling the image autocorrelation by a separable function is that instead of solving the $N^2 \times N^2$ matrix eigenvalue problem of (5.141), only two $N \times N$ matrix eigenvalue problems of (5.146) need to be solved. Since an $N \times N$ matrix eigenvalue problem requires $O(N^3)$ computations, the reduction in dimensionality achieved by the separable model is $O(N^6)/O(N^3) \approx O(N^3)$, which is very significant. Also, the transformation calculations of (5.148) and (5.149) require $2N^3$ operations compared to N^4 operations required for $\Psi^{*T} u$.

Example 5.6

Consider the separable covariance function for a zero mean random field

$$r(m, n; m', n') = \rho^{|m-m'|} \rho^{|n-n'|} \quad (5.150)$$

This gives $\mathcal{R} = \mathbf{R} \otimes \mathbf{R}$, where $\mathbf{R}$ is given by (2.68). The eigenvectors of $\mathbf{R}$ are given by ϕ_k in Example 5.4. Hence $\Psi = \Phi \otimes \Phi$ and the KL transform matrix is $\Phi^{*T} \otimes \Phi^{*T}$. Figure 5.2 shows the basis images of this 8×8 two-dimensional KL transform for $\rho = 0.95$.

Properties of the KL Transform

The KL transform has many desirable properties, which make it optimal in many signal processing applications. Some of these properties are discussed here. For simplicity we assume u has zero mean and a positive definite covariance matrix $\mathbf{R}$.

Decorrelation. The KL transform coefficients $\{v(k), k = 0, \dots, N-1\}$ are uncorrelated and have zero mean, that is,

$$\begin{aligned} E[v(k)] &= 0 \\ E[v(k)v^{*}(l)] &= \lambda_k \delta(k-l) \end{aligned} \quad (5.151)$$

The proof follows directly from (5.133) and (5.135), since

$$E[vv^{*T}] \triangleq \Phi^{*T} E[uu^T] \Phi = \Phi^{*T} \mathbf{R} \Phi = \Lambda = \text{diagonal} \quad (5.152)$$

which implies the latter relation in (5.151). It should be noted that Φ is not a *unique matrix* with respect to this property. There could be many matrices (unitary and nonunitary) that would decorrelate the transformed sequence. For example, a lower triangular matrix Φ could be found that satisfies (5.152).

Example 5.7

The covariance matrix $\mathbf{R}$ of (2.68) is diagonalized by the lower triangular matrix

$$\mathbf{L} \triangleq \begin{bmatrix} 1 & & 0 \\ -\rho & & \\ 0 & -\rho & 1 \end{bmatrix} \Rightarrow \mathbf{L}^T \mathbf{R} \mathbf{L} = \begin{bmatrix} 1-\rho^2 & & 0 \\ & 1-\rho^2 & \\ 0 & & 1-\rho^2 & 1 \\ & & & 1 \end{bmatrix} \triangleq \mathbf{D} \quad (5.153)$$

Hence the transformation $\mathbf{v} = \mathbf{L}^T \mathbf{u}$, will cause the sequence $v(k)$ to be uncorrelated. Comparing with Example 5.4, we see that $\mathbf{L} \neq \Phi$. Moreover, $\mathbf{L}$ is not unitary and the diagonal elements of $\mathbf{D}$ are not the eigenvalues of $\mathbf{R}$.

Basis restriction mean square error. Consider the operations in Fig. 5.16. The vector $\mathbf{u}$ is first transformed to $\mathbf{v}$. The elements of $\mathbf{w}$ are chosen to be the first m elements of $\mathbf{v}$ and zeros elsewhere. Finally, $\mathbf{w}$ is transformed to $\mathbf{z}$. $\mathbf{A}$ and $\mathbf{B}$ are $N \times N$ matrices and $\mathbf{I}_m$ is a matrix with 1s along the first m diagonal terms and zeros elsewhere. Hence

$$w(k) = \begin{cases} v(k), & 0 \leq k \leq m-1 \\ 0, & k \geq m \end{cases} \quad (5.154)$$

Therefore, whereas $\mathbf{u}$ and $\mathbf{v}$ are vectors in an N -dimensional vector space, $\mathbf{w}$ is a vector restricted to an $m \leq N$ -dimensional subspace. The average mean square error between the sequences $u(n)$ and $z(n)$ is defined as

$$J_m \triangleq \frac{1}{N} E \left(\sum_{n=0}^{N-1} |u(n) - z(n)|^2 \right) = \frac{1}{N} \text{Tr} \{ E \{ (\mathbf{u} - \mathbf{z})(\mathbf{u} - \mathbf{z})^* \} \} \quad (5.155)$$

This quantity is called the *basis restriction error*. It is desired to find the matrices $\mathbf{A}$ and $\mathbf{B}$ such that J_m is minimized for each and every value of $m \in [1, N]$. This minimum is achieved by the KL transform of $\mathbf{u}$.

Theorem 5.1. The error J_m in (5.155) is minimum when

$$\mathbf{A} = \Phi^{*T}, \quad \mathbf{B} = \Phi, \quad \mathbf{A}\mathbf{B} = \mathbf{I} \quad (5.156)$$

where the columns of Φ are arranged according to the decreasing order of the eigenvalues of $\mathbf{R}$.

Proof. From Fig. 5.16, we have

$$\mathbf{v} = \mathbf{A}\mathbf{u}, \quad \mathbf{w} = \mathbf{I}_m \mathbf{v}, \quad \text{and} \quad \mathbf{z} = \mathbf{B}\mathbf{w} \quad (5.157)$$

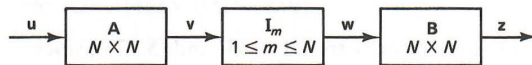


Figure 5.16 KL transform basis restriction.

Using these we can rewrite (5.155) as

$$J_m = \frac{1}{N} \text{Tr}[(\mathbf{I} - \mathbf{B}\mathbf{I}_m\mathbf{A})\mathbf{R}(\mathbf{I} - \mathbf{B}\mathbf{I}_m\mathbf{A})^*]^T]$$

To minimize J_m we first differentiate it with respect to the elements of $\mathbf{A}$ and set the result to zero [see Problem 2.15 for the differentiation rules]. This gives

$$\mathbf{I}_m\mathbf{B}^T(\mathbf{I} - \mathbf{B}\mathbf{I}_m\mathbf{A})^*\mathbf{R} = \mathbf{0} \quad (5.158)$$

which yields

$$J_m = \frac{1}{N} \text{Tr}[(\mathbf{I} - \mathbf{B}\mathbf{I}_m\mathbf{A})\mathbf{R}] \quad (5.159)$$

$$\mathbf{I}_m\mathbf{B}^{*T} = \mathbf{I}_m\mathbf{B}^{*T}\mathbf{B}\mathbf{I}_m\mathbf{A} \quad (5.160)$$

At $m = N$, the minimum value of J_N must be zero, which requires

$$\mathbf{I} - \mathbf{B}\mathbf{A} = \mathbf{0} \quad \text{or} \quad \mathbf{B} = \mathbf{A}^{-1} \quad (5.161)$$

Using this in (5.160) and rearranging terms, we obtain

$$\mathbf{I}_m\mathbf{B}^{*T}\mathbf{B} = \mathbf{I}_m\mathbf{B}^{*T}\mathbf{B}\mathbf{I}_m, \quad 1 \leq m \leq N \quad (5.162)$$

For (5.162) to be true for every m , it is necessary that $\mathbf{B}^{*T}\mathbf{B}$ be diagonal. Since $\mathbf{B} = \mathbf{A}^{-1}$, it is easy to see that (5.160) remains invariant if $\mathbf{B}$ is replaced by $\mathbf{D}\mathbf{B}$ or $\mathbf{B}\mathbf{D}$, where $\mathbf{D}$ is a diagonal matrix. Hence, without loss of generality we can normalize $\mathbf{B}$ so that $\mathbf{B}^{*T}\mathbf{B} = \mathbf{I}$, that is, $\mathbf{B}$ is a unitary matrix. Therefore, $\mathbf{A}$ is also unitary and $\mathbf{B} = \mathbf{A}^{*T}$. This gives

$$J_m = \frac{1}{N} \text{Tr}[(\mathbf{I} - \mathbf{A}^{*T}\mathbf{I}_m\mathbf{A})\mathbf{R}] = \frac{1}{N} \text{Tr}(\mathbf{R} - \mathbf{I}_m\mathbf{A}\mathbf{R}\mathbf{A}^{*T}) \quad (5.163)$$

Since $\mathbf{R}$ is fixed, J_m is minimized if the quantity

$$\bar{J}_m \triangleq \text{Tr}(\mathbf{I}_m\mathbf{A}\mathbf{R}\mathbf{A}^{*T}) = \sum_{k=0}^{m-1} \mathbf{a}_k^T \mathbf{R} \mathbf{a}_k^* \quad (5.164)$$

is maximized where $\mathbf{a}_k^T$ is the k th row of $\mathbf{A}$. Since $\mathbf{A}$ is unitary,

$$\mathbf{a}_k^T \mathbf{a}_k^* = 1 \quad (5.165)$$

To maximize $\bar{J}_m$ subject to (5.165), we form the Lagrangian

$$\bar{J}_m = \sum_{k=0}^{m-1} \mathbf{a}_k^T \mathbf{R} \mathbf{a}_k^* + \sum_{k=0}^{m-1} \lambda_k (1 - \mathbf{a}_k^T \mathbf{a}_k^*) \quad (5.166)$$

and differentiate it with respect to $\mathbf{a}_j$. The result gives a necessary condition

$$\mathbf{R} \mathbf{a}_j^* = \lambda_j \mathbf{a}_j^* \quad (5.167)$$

where $\mathbf{a}_j^*$ are orthonormalized eigenvectors of $\mathbf{R}$. This yields

$$\bar{J}_m = \sum_{k=0}^{m-1} \lambda_k \quad (5.168)$$

which is maximized if $\{\mathbf{a}_j^*, 0 \leq j \leq m-1\}$ correspond to the largest m eigenvalues of $\mathbf{R}$. Because $\tilde{J}_m$ must be maximized for every m , it is necessary to arrange $\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{N-1}$. Then $\mathbf{a}_j^T$, the rows of $\mathbf{A}$, are the conjugate transpose of the eigenvectors of $\mathbf{R}$, that is, $\mathbf{A}$ is the KL transform of $\mathbf{u}$.

Distribution of variances. Among all the unitary transformations $\mathbf{v} = \mathbf{A}\mathbf{u}$, the KL transform Φ^{*T} packs the maximum average energy in $m \leq N$ samples of $\mathbf{v}$. Define

$$\sigma_k^2 \triangleq E[|v(k)|^2], \quad \sigma_0^2 \geq \sigma_1^2 \geq \dots \geq \sigma_{N-1}^2$$

$$S_m(\mathbf{A}) \triangleq \sum_{k=0}^{m-1} \sigma_k^2 \quad (5.169)$$

Then for any fixed $m \in [1, N]$

$$S_m(\Phi^{*T}) \geq S_m(\mathbf{A}) \quad (5.170)$$

Proof. Note that

$$\begin{aligned} S_m(\mathbf{A}) &= \sum_{k=0}^{m-1} (\mathbf{A}\mathbf{R}\mathbf{A}^{*T})_{k,k} \\ &= \text{Tr}(\mathbf{I}_m \mathbf{A}^{*T} \mathbf{R} \mathbf{A}) \\ &= \tilde{J}_m \end{aligned}$$

which, we know from the last property [see (5.164)], is maximized when $\mathbf{A}$ is the KL transform. Since $\sigma_k^2 = \lambda_k$ when $\mathbf{A} = \Phi^{*T}$, from (5.168)

$$\sum_{k=0}^{m-1} \lambda_k \geq \sum_{k=0}^{m-1} \sigma_k^2, \quad 1 \leq m \leq N \quad (5.171)$$

Threshold representation. The KL transform also minimizes $E[m]$, the expected number of transform coefficients required, so that their energy just exceeds a prescribed threshold (see Problem 5.26 and [33]).

A fast KL transform. In application of the KL transform to images, there are dimensionality difficulties. The KL transform depends on the statistics as well as the size of the image and, in general, the basis vectors are not known analytically. After the transform matrix has been computed, the operations for performing the transformation are quite large for images.

It has been shown that certain statistical image models yield a fast KL transform algorithm as an alternative to the conventional KL transform for images. It is based on a stochastic decomposition of an image as a sum of two random sequences. *The first random sequence is such that its KL transform is a fast transform and the second sequence, called the boundary response, depends only on information at the boundary points of the image.* For details see Sections 6.5 and 6.9.

The rate-distortion function. Suppose a random vector $\mathbf{u}$ is unitary transformed to $\mathbf{v}$ and transmitted over a communication channel (Fig. 5.17). Let $\mathbf{v}'$ and $\mathbf{u}'$

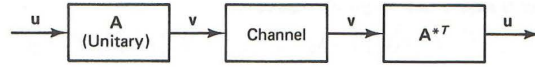


Figure 5.17 Unitary transform data transmission. Each element of $\mathbf{v}$ is coded independently.

be the reproduced values of $\mathbf{v}$ and $\mathbf{u}$, respectively. Further, assume that $\mathbf{u}$, $\mathbf{v}$, $\mathbf{v}'$, and $\mathbf{u}'$ are Gaussian. The average distortion in $\mathbf{u}$ is

$$D = \frac{1}{N} E[(\mathbf{u} - \mathbf{u}')^* T (\mathbf{u} - \mathbf{u}')] \quad (5.172)$$

Since $\mathbf{A}$ is unitary and $\mathbf{u} = \mathbf{A}^{*T} \mathbf{v}$ and $\mathbf{u}' = \mathbf{A}^{*T} \mathbf{v}'$, we have

$$\begin{aligned} D &= \frac{1}{N} E[(\mathbf{v} - \mathbf{v}')^* T \mathbf{A} \mathbf{A}^{*T} (\mathbf{v} - \mathbf{v}')] \\ &= \frac{1}{N} E[(\mathbf{v} - \mathbf{v}')^* T (\mathbf{v} - \mathbf{v}')] = \frac{1}{N} E(\delta \mathbf{v}^{*T} \delta \mathbf{v}) \end{aligned} \quad (5.173)$$

where $\delta \mathbf{v} = \mathbf{v} - \mathbf{v}'$ represents the error in the reproduction of $\mathbf{v}$. From the preceding, D is invariant under all unitary transformations. The rate-distortion function is now obtained, following Section 2.13, as

$$R = \frac{1}{N} \sum_{k=0}^{N-1} \max \left[0, \frac{1}{2} \log_2 \frac{\sigma_k^2}{\theta} \right] \quad (5.174)$$

$$D = \frac{1}{N} \sum_{k=0}^{N-1} \min[\theta, \sigma_k^2] \quad (5.175)$$

where

$$\sigma_k^2 = E[|v(k)|^2] = [\mathbf{A} \mathbf{R} \mathbf{A}^{*T}]_{k,k} \quad (5.176)$$

depend on the transform $\mathbf{A}$. Therefore, the rate

$$R = R(\mathbf{A}) \quad (5.177)$$

also depends on $\mathbf{A}$. For each fixed D , the KL transform achieves the minimum rate among all unitary transforms, that is,

$$R(\Phi^{*T}) \leq R(\mathbf{A}) \quad (5.178)$$

This property is discussed further in Chapter 11, on transform coding.

Example 5.8

Consider a 2×1 vector $\mathbf{u}$, whose covariance matrix is

$$\mathbf{R} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \quad |\rho| < 1$$

The KL transform is

$$\Phi^{*T} = \Phi = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

The transformation $\mathbf{v} = \Phi \mathbf{u}$ gives

$$E\{[v(0)]^2\} = \lambda_0 = 1 + \rho, \quad E\{[v(1)]^2\} = 1 - \rho$$

$$R(\Phi) = \frac{1}{2} \left[\max \left(0, \frac{1}{2} \log \frac{1+\rho}{\theta} \right) + \max \left(0, \frac{1}{2} \log \frac{1-\rho}{\theta} \right) \right]$$

Compare this with the case when $\mathbf{A} = \mathbf{I}$ (that is, $\mathbf{u}$ is transmitted), which gives $\sigma_0^2 = \sigma_1^2 = 1$, and

$$R(\mathbf{I}) = \frac{1}{4} [-2 \log \theta], \quad 0 < \theta < 1$$

Suppose we let θ be small, say $\theta < 1 - |\rho|$. Then it is easy to show that

$$R(\Phi) < R(\mathbf{I})$$

This means for a fixed level of distortion, the number of bits required to transmit the KLT sequence would be less than those required for transmission of the original sequence.

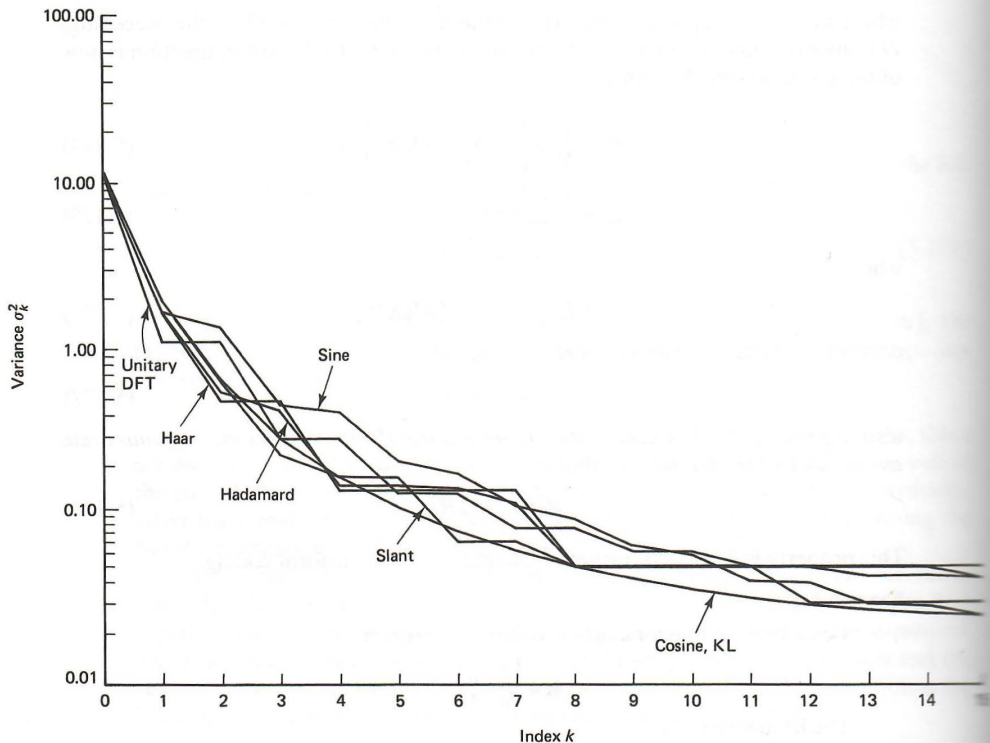


Figure 5.18 Distribution of variances of the transform coefficients (in decreasing order) of a stationary Markov sequence with $N = 16$, $\rho = 0.95$ (see Example 5.9).

TABLE 5.2 Variances σ_k^2 of Transform Coefficients of a Stationary Markov Sequence with $\rho = 0.95$ and $N = 16$. See Example 5.9.

Transform ↓ k	KL	Cosine	Sine	Unitary DFT	Hadamard	Haar	Slant
0	12.442	12.406	11.169	12.406	12.406	12.406	12.406
1	1.946	1.943	1.688	1.100	1.644	1.644	1.904
2	0.615	0.648	1.352	0.292	0.544	0.487	0.641
3	0.292	0.295	0.421	0.139	0.431	0.487	0.233
4	0.171	0.174	0.463	0.086	0.153	0.144	0.173
5	0.114	0.114	0.181	0.062	0.152	0.144	0.172
6	0.082	0.083	0.216	0.051	0.149	0.144	0.072
7	0.063	0.063	0.098	0.045	0.121	0.144	0.072
8	0.051	0.051	0.116	0.043	0.051	0.050	0.051
9	0.043	0.043	0.060	0.045	0.051	0.050	0.051
10	0.037	0.037	0.067	0.051	0.051	0.050	0.051
11	0.033	0.033	0.040	0.062	0.051	0.050	0.051
12	0.030	0.030	0.042	0.086	0.051	0.050	0.031
13	0.028	0.028	0.031	0.139	0.051	0.050	0.031
14	0.027	0.027	0.029	0.292	0.050	0.050	0.031
15	0.026	0.026	0.026	1.100	0.043	0.050	0.031

Example 5.9 (Comparison Among Unitary Transforms for a Markov Sequence)

Consider a first-order zero mean stationary Markov sequence of length N whose covariance matrix $\mathbf{R}$ is given by (2.68) with $\rho = 0.95$. Figure 5.18 shows the distribution of variances σ_k^2 of the transform coefficients (in decreasing order) for different transforms. Table 5.2 lists σ_k^2 for the various transforms.

Define the normalized basis restriction error as

$$J_m = \frac{\sum_{k=m}^{N-1} \sigma_k^2}{\sum_{k=0}^{N-1} \sigma_k^2}, \quad m = 0, \dots, N-1 \quad (5.179)$$

where σ_k^2 have been arranged in decreasing order.

Figure 5.19 shows J_m versus m for the various transforms. It is seen that the cosine transform performance is indistinguishable from that of the KL transform for $\rho = 0.95$. In general it seems possible to find a fast sinusoidal transform (that is, a transform consisting of sine or cosine functions) as a good substitute for the KL transform for different values of ρ as well as for higher-order stationary random sequences (see Section 5.12).

The mean square performance of the various transforms also depends on the dimension N of the transform. Such comparisons are made in Section 5.12.

Example 5.10 (Performance of Transforms on Images)

The mean square error test of the last example can be extended to actual images. Consider an $N \times N$ image $u(m, n)$ from which its mean is subtracted out to make it zero mean. The transform coefficient variances are estimated as

$$\sigma^2(k, l) = E[|v(k, l)|^2] \cong |v(k, l)|^2$$

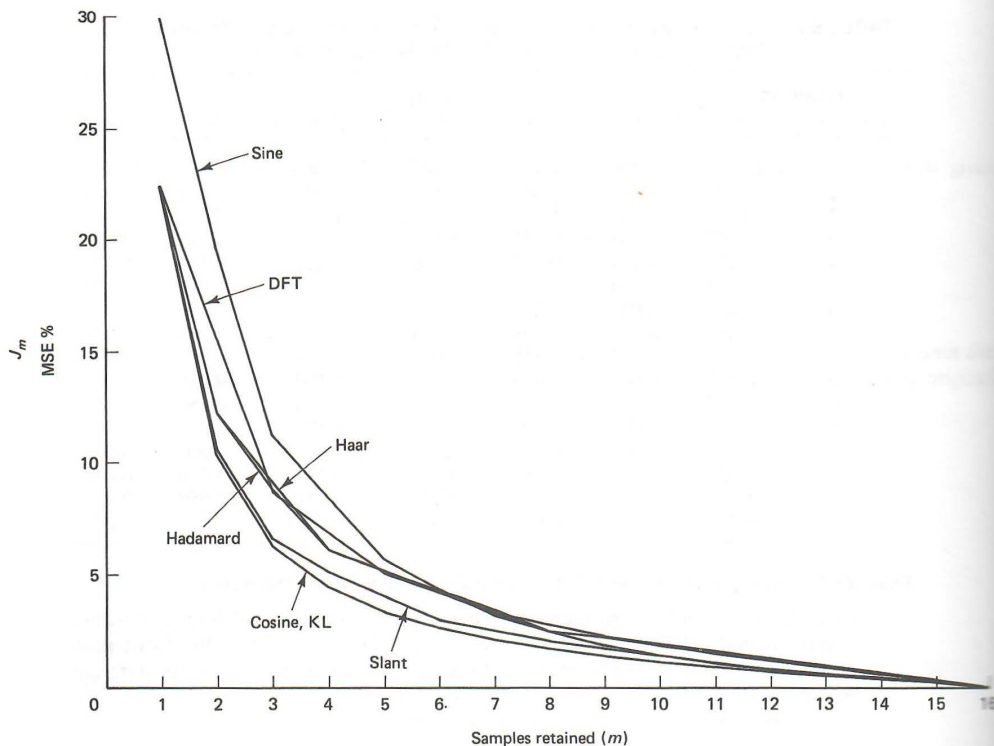


Figure 5.19 Performance of different unitary transforms with respect to basis restriction errors (J_m) versus the number of basis (m) for a stationary Markov sequence with $N = 16$, $\rho = 0.95$.

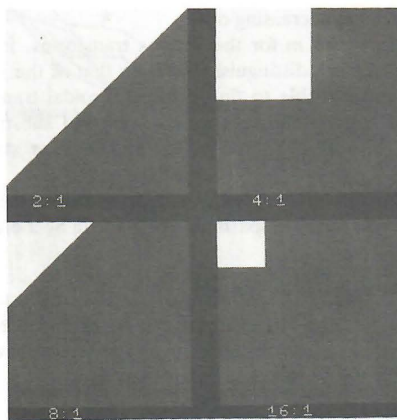


Figure 5.20 Zonal filters for 2:1, 4:1, 8:1, 16:1 sample reduction. White areas are passbands, dark areas are stopbands.

The image transform is filtered by a *zonal mask* (Fig. 5.20) such that only a fraction of the transform coefficients are retained and the remaining ones are set to zero. Define the normalized mean square error

$$J_s \triangleq \frac{\sum_{k,l \in \text{stopband}} |v_{k,l}|^2}{\sum_{k,l=0}^{N-1} |v_{k,l}|^2} = \frac{\text{energy in stopband}}{\text{total energy}}$$

Figure 5.21 shows an original image and the image obtained after cosine transform zonal filtering to achieve various sample reduction ratios. Figure 5.22 shows the zonal filtered images for different transforms at a 4:1 sample reduction ratio. Figure 5.23 shows the mean square error versus sample reduction ratio for different transforms. Again we find the cosine transform to have the best performance.



Figure 5.21 Basis restriction zonal filtered images in cosine transform domain.



(a) Cosine;



(b) sine;



(c) unitary DFT;



(d) Hadamard;



(e) Haar;



(f) Slant.

Figure 5.22 Basis restriction zonal filtering using different transforms with 4:1 sample reduction.

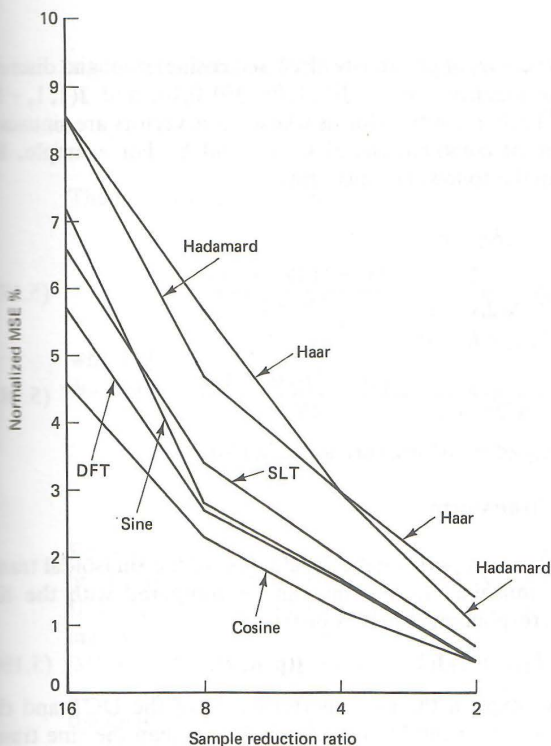


Figure 5.23 Performance comparison of different transforms with respect to basis restriction zonal filtering for 256×256 images.

5.12 A SINUSOIDAL FAMILY OF UNITARY TRANSFORMS

This is a class of complete orthonormal sets of eigenvectors generated by the parametric family of matrices whose structure is similar to that of $\mathbf{R}^{-1}$ [see (5.96)],

$$\mathbf{J} = \mathbf{J}(k_1, k_2, k_3) = \begin{bmatrix} 1 - k_1 \alpha & -\alpha & k_3 \alpha \\ -\alpha & 1 & 0 \\ k_3 \alpha & 0 & 1 - k_2 \alpha \end{bmatrix} \quad (5.180)$$

In fact, for $k_1 = k_2 = \rho$, $k_3 = 0$, $\beta^2 = (1 - \rho^2)/(1 + \rho^2)$, and $\alpha = \rho/(1 + \rho^2)$, we have

$$\mathbf{J}(\rho, \rho, 0) = \beta^2 \mathbf{R}^{-1} \quad (5.181)$$

Since $\mathbf{R}$ and $\beta^2 \mathbf{R}^{-1}$ have an identical set of eigenvectors, the KL transform associated with $\mathbf{R}$ can be determined from the eigenvectors of $\mathbf{J}(\rho, \rho, 0)$. Similarly, it

can be shown that the basis vectors of previously discussed cosine, sine, and discrete Fourier transforms are the eigenvectors of $\mathbf{J}(1, 1, 0)$, $\mathbf{J}(0, 0, 0)$, and $\mathbf{J}(1, 1, -1)$, respectively. In fact, several other fast transforms whose basis vectors are sinusoids can be generated for different combinations of k_1 , k_2 , and k_3 . For example, for $0 \leq m, k \leq N - 1$, we obtain the following transforms:

1. Odd sine - 1: $k_1 = k_3 = 0, k_2 = 1$

$$\phi_m(k) = \frac{2}{\sqrt{2N+1}} \sin \frac{(k+1)(2m+1)\pi}{2N+1} \quad (5.182)$$

2. Odd cosine - 1: $k_1 = 1, k_2 = k_3 = 0$

$$\phi_m(k) = \frac{2}{\sqrt{2N+1}} \cos \frac{(2k+1)(2m+1)\pi}{2(2N+1)} \quad (5.183)$$

Other members of this family of transforms are given in [34].

Approximation to the KL Transform

The $\mathbf{J}$ matrices play a useful role in performance evaluation of the sinusoidal transforms. For example, two sinusoidal transforms can be compared with the KL transform by comparing corresponding $\mathbf{J}$ -matrix distances

$$\Delta(k_1, k_2, k_3) \triangleq \|\mathbf{J}(k_1, k_2, k_3) - \mathbf{J}(\rho, \rho, 0)\|^2 \quad (5.184)$$

This measure can also explain the close performance of the DCT and the KLT. Further, it can be shown that the DCT performs better than the sine transform for $0.5 \leq \rho \leq 1$ and the sine transform performs better than the cosine for other values of ρ . The $\mathbf{J}$ matrices are also useful in finding a fast sinusoidal transform approximation to the KL transform of an arbitrary random sequence whose covariance matrix is $\mathbf{A}$. If $\mathbf{A}$ commutes with a $\mathbf{J}$ matrix, that is, $\mathbf{A}\mathbf{J} = \mathbf{J}\mathbf{A}$, then they will have an identical set of eigenvectors. The best fast sinusoidal transform may be chosen as the one whose corresponding $\mathbf{J}$ matrix minimizes the commuting distance $\|\mathbf{A}\mathbf{J} - \mathbf{J}\mathbf{A}\|^2$. Other uses of the $\mathbf{J}$ matrices are (1) finding fast algorithms for inversion of banded Toeplitz matrices, (2) efficient calculation of transform coefficient variances, which are needed in transform domain processing algorithms, and (3) establishing certain useful asymptotic properties of these transforms. For details see [34].

5.13 OUTER PRODUCT EXPANSION AND SINGULAR VALUE DECOMPOSITION

In the foregoing transform theory, we considered an $N \times M$ image $\mathbf{U}$ to be a vector in an NM -dimensional vector space. However, it is possible to represent any such image in an r -dimensional subspace where r is the rank of the matrix $\mathbf{U}$.

Let the image be real and $M \leq N$. The matrices $\mathbf{U}\mathbf{U}^T$ and $\mathbf{U}^T\mathbf{U}$ are non-negative, symmetric and have the identical eigenvalues, $\{\lambda_m\}$. Since $M \leq N$, there are at most $r \leq M$ nonzero eigenvalues. It is possible to find r orthogonal, $M \times 1$

eigenvectors $\{\phi_m\}$ of $U^T U$ and r orthogonal $N \times 1$ eigenvectors $\{\psi_m\}$ of UU^T , that is,

$$U^T U \phi_m = \lambda_m \phi_m, \quad m = 1, \dots, r \quad (5.185)$$

$$UU^T \psi_m = \lambda_m \psi_m, \quad m = 1, \dots, r \quad (5.186)$$

The matrix U has the representation

$$U = \Psi \Lambda^{1/2} \Phi^T \quad (5.187)$$

$$= \sum_{m=1}^r \sqrt{\lambda_m} \psi_m \phi_m^T \quad (5.188)$$

where Ψ and Φ are $N \times r$ and $M \times r$ matrices whose m th columns are the vectors ψ_m and ϕ_m , respectively, and $\Lambda^{1/2}$ is an $r \times r$ diagonal matrix, defined as

$$\Lambda^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_r} \end{bmatrix} \quad (5.189)$$

Equation (5.188) is called the *spectral representation*, the *outer product expansion*, or the *singular value decomposition* (SVD) of U . The nonzero eigenvalues (of $U^T U$), λ_m , are also called the singular values of U . If $r \ll M$, then the image containing NM samples can be represented by $(M+N)r$ samples of the vectors $\{\lambda_m^{1/4} \psi_m, \lambda_m^{1/4} \phi_m; m = 1, \dots, r\}$.

Since Ψ and Φ have orthogonal columns, from (5.187) the *SVD transform* of the image U is defined as

$$\Lambda^{1/2} = \Psi^T U \Phi \quad (5.190)$$

which is a separable transform that diagonalizes the given image. The proof of (5.188) is outlined in Problem 5.31.

Properties of the SVD Transform

1. Once $\phi_m, m = 1, \dots, r$ are known, the eigenvectors ψ_m can be determined as

$$\psi_m \triangleq \frac{1}{\sqrt{\lambda_m}} U \phi_m, \quad m = 1, \dots, r \quad (5.191)$$

It can be shown that ψ_m are orthonormal eigenvectors of UU^T if ϕ_m are the orthonormal eigenvectors of $U^T U$.

2. The SVD transform as defined by (5.190) is not a unitary transform. This is because Ψ and Φ are rectangular matrices. However, we can include in Φ and Ψ additional orthogonal eigenvectors ϕ_m and ψ_m , which satisfy $U \phi_m = 0, m = r+1, \dots, M$ and $U^T \psi_m = 0, m = r+1, \dots, N$ such that these matrices are unitary and the *unitary SVD transform* is

$$\begin{bmatrix} \Lambda^{1/2} \\ 0 \end{bmatrix} = \Psi^T U \Phi \quad (5.192)$$

3. The image U_k generated by the partial sum

$$U_k \triangleq \sum_{m=1}^k \sqrt{\lambda_m} \psi_m \phi_m^T, \quad k \leq r \quad (5.193)$$

is the best least squares rank- k approximation of U if λ_m are in decreasing order of magnitude. For any $k \leq r$, the least squares error

$$\epsilon_k^2 = \sum_{m=1}^M \sum_{n=1}^N |u(m, n) - u_k(m, n)|^2, \quad k = 1, 2, \dots, r \quad (5.194)$$

reduces to

$$\epsilon_k^2 = \sum_{m=k+1}^r \lambda_m \quad (5.195)$$

Let $L \triangleq NM$. Note that we can always write a two-dimensional unitary transform representation as an outer product expansion in an L -dimensional space, namely,

$$U = \sum_{l=1}^L w_l \mathbf{a}_l \mathbf{b}_l^T \quad (5.196)$$

where w_l are scalars and $\mathbf{a}_l$ and $\mathbf{b}_l$ are sequences of orthogonal basis vectors of dimensions $N \times 1$ and $M \times 1$, respectively. The least squares error between U and any partial sum

$$\hat{U}_k \triangleq \sum_{l=1}^k w_l \mathbf{a}_l \mathbf{b}_l^T \quad (5.197)$$

is minimized for any $k \in [1, L]$ when the above expansion coincides with (5.193), that is, when $\hat{U}_k = U_k$.

This means *the energy concentrated in the transform coefficients $w_l, l = 1, \dots, k$ is maximized by the SVD transform for the given image*. Recall that the KL transform, maximizes *the average energy in a given number of transform coefficients, the average being taken over the ensemble for which the autocorrelation function is defined*. Hence, on an image-to-image basis, the SVD transform will concentrate more energy in the same number of coefficients. But the SVD has to be calculated for *each* image. On the other hand the KL transform needs to be calculated *only* once for the whole image ensemble. Therefore, while one may be able to find a reasonable fast transform approximation of the KL transform, no such fast transform substitute for the SVD is expected to exist.

Although applicable in image restoration and image data compression problems, the usefulness of SVD in such image processing problems is severely limited because of large computational effort required for calculating the eigenvalues and eigenvectors of large image matrices. However, the SVD is a fundamental result in matrix theory that is useful in finding the generalized inverse of singular matrices and in the analysis of several image processing problems.

Example 5.11

Let

$$U = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{bmatrix}$$

The eigenvalues of $\mathbf{U}^T \mathbf{U}$ are found to be $\lambda_1 = 18.06$, $\lambda_2 = 1.94$, which give $r = 2$, and the SVD transform of $\mathbf{U}$ is

$$\mathbf{\Lambda}^{1/2} = \begin{bmatrix} 4.25 & 0 \\ 0 & 1.39 \end{bmatrix}$$

The eigenvectors are found to be

$$\Phi_1 = \begin{bmatrix} 0.5019 \\ 0.8649 \end{bmatrix}, \quad \Phi_2 = \begin{bmatrix} 0.8649 \\ -0.5019 \end{bmatrix}$$

(continued on page 180)

TABLE 5.3 Summary of Image Transforms

DFT/unitary DFT	Fast transform, most useful in digital signal processing, convolution, digital filtering, analysis of circulant and Toeplitz systems. Requires complex arithmetic. Has very good energy compaction for images.
Cosine	Fast transform, requires real operations, near optimal substitute for the KL transform of highly correlated images. Useful in designing transform coders and Wiener filters for images. Has excellent energy compaction for images.
Sine	About twice as fast as the fast cosine transform, symmetric, requires real operations; yields fast KL transform algorithm which yields recursive block processing algorithms, for coding, filtering, and so on; useful in estimating performance bounds of many image processing problems. Energy compaction for images is very good.
Hadamard	Faster than sinusoidal transforms, since no multiplications are required; useful in digital hardware implementations of image processing algorithms. Easy to simulate but difficult to analyze. Applications in image data compression, filtering, and design of codes. Has good energy compaction for images.
Haar	<u>Very fast transform</u> . Useful in feature extraction, image coding, and image analysis problems. Energy compaction is fair.
Slant	Fast transform. Has "image-like basis"; useful in image coding. Has very good energy compaction for images.
Karhunen-Loeve	Is optimal in many ways; has no fast algorithm; useful in performance evaluation and for finding performance bounds. Useful for small size vectors e.g., color multispectral or other feature vectors. Has the best energy compaction in the mean square sense over an ensemble.
Fast KL	Useful for designing fast, recursive-block processing techniques, including adaptive techniques. Its performance is better than independent block-by-block processing techniques.
Sinusoidal transforms	Many members have fast implementation, useful in finding practical substitutes for the KL transform, analysis of Toeplitz systems, mathematical modeling of signals. Energy compaction for the optimum-fast transform is excellent.
SVD transform	Best energy-packing efficiency for any given image. Varies drastically from image to image; has no fast algorithm or a reasonable fast transform substitute; useful in design of separable FIR filters, finding least squares and minimum norm solutions of linear equations, finding rank of large matrices, and so on. Potential image processing applications are in image restoration, power spectrum estimation and data compression.

From above ψ_1 is obtained via (5.191) to yield

$$\mathbf{U}_1 = \sqrt{\lambda_1} \psi_1 \phi_1^T = \begin{bmatrix} 1.120 & 1.94 \\ 0.935 & 1.62 \\ 1.549 & 2.70 \end{bmatrix}$$

as the best least squares rank-1 approximation of $\mathbf{U}$. Let us compare this with the two dimensional cosine transform $\mathbf{U}$, which is given by

$$\mathbf{V} = \mathbf{C}_3 \mathbf{U} \mathbf{C}_2^T = \frac{1}{\sqrt{12}} \begin{bmatrix} \sqrt{2} & \sqrt{2} & \sqrt{2} \\ \sqrt{3} & 0 & -\sqrt{3} \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{\sqrt{12}} \begin{bmatrix} 10\sqrt{2} & -2\sqrt{2} \\ -\sqrt{3} & \sqrt{3} \\ -1 & -5 \end{bmatrix}$$

It is easy to see that $\sum_{k,l} v^2(k, l) = \lambda_1 + \lambda_2$. The energy concentrated in the K samples of SVD, $\sum_{m=1}^K \lambda_m$, $K = 1, 2$, is greater than the energy concentrated in any K samples of the cosine transform coefficients (show!).

5.14 SUMMARY

In this chapter we have studied the theory of unitary transforms and their properties. Several unitary transforms, DFT, cosine, sine, Hadamard, Haar, Slant, KL, sinusoidal family, fast KL, and SVD, were discussed. Table 5.3 summarizes the various transforms and their applications.

PROBLEMS

- 5.1 For given P, Q show that the error σ_e^2 of (5.8) is minimized when the series coefficients $v(k, l)$ are given by (5.3). Also show that the basis images must form a complete set for σ_e^2 to be zero for $P = Q = N$.
- 5.2 (*Fast transforms and Kronecker separability*) From (5.23) we see that the number of operations in implementing the matrix-vector product is reduced from $O(N^4)$ to $O(N^3)$ if $\mathcal{A}$ is a Kronecker product. Apply this idea inductively to show that if $\mathcal{A}$ is $M \times M$ and

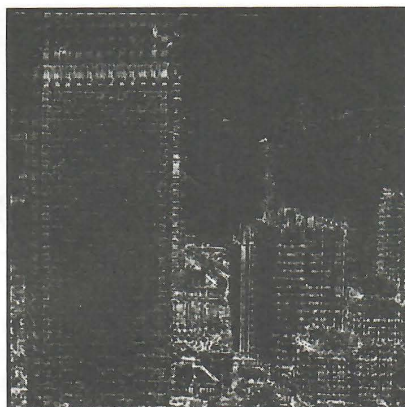
$$\mathcal{A} = \mathbf{A}_1 \otimes \mathbf{A}_2 \otimes \dots \otimes \mathbf{A}_m$$

where $\mathbf{A}_k$ is $n_k \times n_k$, $M = \prod_{k=1}^m n_k$, then the transformation of (5.23) can be implemented in $O(M \sum_{k=1}^m n_k)$, which equals $nM \log_n M$ if $n_k = n$. Many fast algorithms for unitary matrices can be given this interpretation which was suggested by Good [9]. Transforms possessing this property are sometimes called *Good transforms*.

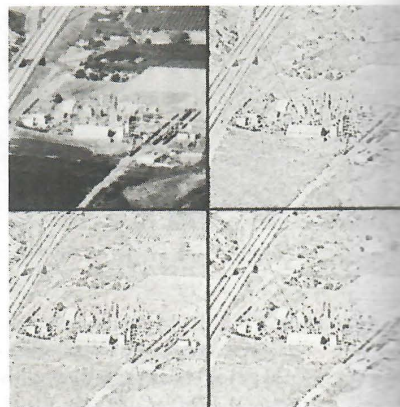
- 5.3 For the 2×2 transform $\mathbf{A}$ and the image $\mathbf{U}$

$$\mathbf{A} = \frac{1}{2} \begin{bmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$$

calculate the transformed image $\mathbf{V}$ and the basis images.



cepstrum of the building image



generalized cepstra

- (a) original
(b) DFT
(c) DCT
(d) Hadamard transform

Figure 7.37 Generalized cepstra:

called the generalized homomorphic transform, $\mathcal{H}$, of the image $u(m, n)$. The generalized homomorphic linear filter performs zero-memory operations on the $\mathcal{H}$ -transform of the image followed by inverse $\mathcal{H}$ -transform, as shown in Fig. 7.36. Examples of cepstra are given in Fig. 7.37. The homomorphic transformation reduces the dynamic range of the image in the transform domain and increases it in the cepstral domain.

7.6 MULTISPECTRAL IMAGE ENHANCEMENT

In multispectral imaging there is a sequence of I images $U_i(m, n)$, $i = 1, 2, \dots, I$, where the number I is typically between 2 and 12. It is desired to combine these images to generate a single or a few display images that are representative of their features. There are three common methods of enhancing such images.

Intensity Ratios

Define the ratios

$$R_{i,j}(m, n) \triangleq \frac{u_i(m, n)}{u_j(m, n)}, \quad i \neq j \quad (7.53)$$

where $u_i(m, n)$ represents the intensity and is assumed to be positive. This method gives $I^2 - I$ combinations for the ratios, the few most suitable of which are chosen by visual inspection. Sometimes the ratios are defined with respect to the average image $(1/I) \sum_{i=1}^I u_i(m, n)$ to reduce the number of combinations.

Log-Ratios

Taking the logarithm on both sides of (7.53), we get

$$L_{i,j} \triangleq \log R_{i,j} = \log u_i(m, n) - \log u_j(m, n) \quad (7.54)$$

The log-ratio $L_{i,j}$ gives a better display when the dynamic range of $R_{i,j}$ is very large, which can occur if the spectral features at a spatial location are quite different.

Principal Components

For each (m, n) define the $I \times 1$ vector

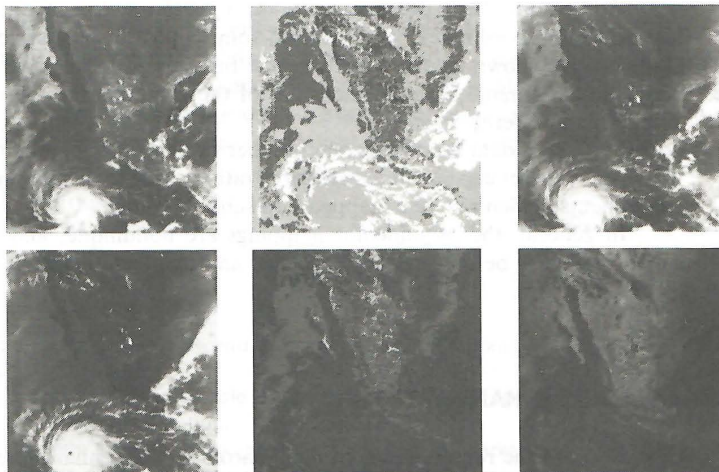
$$\mathbf{u}(m, n) = \begin{bmatrix} u_1(m, n) \\ u_2(m, n) \\ \vdots \\ u_I(m, n) \end{bmatrix} \quad (7.55)$$

The $I \times I$ KL transform of $\mathbf{u}(m, n)$, denoted by Φ , is determined from the autocorrelation matrix of the ensemble of vectors $\{\mathbf{u}_i(m, n), i = 1, \dots, I\}$. The rows of Φ , which are eigenvectors of the autocorrelation matrix, are arranged in decreasing order of their associated eigenvalues. Then for any $I_0 \leq I$, the images $v_i(m, n)$, $i = 1, \dots, I_0$ obtained from the KL transformed vector

$$\mathbf{v}(m, n) = \Phi \mathbf{u}(m, n) \quad (7.56)$$

are the first I_0 principal components of the multispectral images.

Figure 7.38 contains examples of multispectral image enhancement.



Originals (top), visual band: (bottom), I.R.

Log Ratios

Principal Components

Figure 7.38 Multispectral image enhancement. The clouds and land have been separated.