

Image and Video Processing in the Compressed Domain

Jayanta Mukhopadhyay

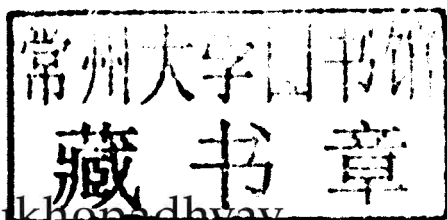


CRC Press
Taylor & Francis Group

A CHAPMAN & HALL BOOK

Image and Video Processing in the Compressed Domain

Jayanta Mukhopadhyay



CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

A CHAPMAN & HALL BOOK

MATLAB® is a trademark of The MathWorks, Inc. and is used with permission. The MathWorks does not warrant the accuracy of the text or exercises in this book. This book's use or discussion of MATLAB® software or related products does not constitute endorsement or sponsorship by The MathWorks of a particular pedagogical approach or particular use of the MATLAB® software.

Chapman & Hall/CRC
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2011 by Taylor and Francis Group, LLC
Chapman & Hall/CRC is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number: 978-1-4398-2935-6 (Hardback)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the CRC Press Web site at
<http://www.crcpress.com>

Image and Video Processing in the Compressed Domain

Dedicated to my parents.

Preface

With more and more images and videos being available in the compressed format, researchers have started taking interest in the aspect of designing algorithms for different image operations directly in their domains of representation. This would not only avoid the inverse and forward transformation steps and potentially make the computation faster but also would keep the buffer requirement less as the storage required for processing the compressed stream is less than that with its uncompressed representation.

This book attempts to comprehensively treat this topic of interest and deals with the fundamentals and properties of various image transforms used in image and video compression. Subsequently, their application in designing image and video processing algorithms in the compressed domain are discussed. To provide better understanding of the domain of research, different image and video compression techniques are briefly covered in the first chapter. In particular, discrete cosine transform (DCT) - based compression algorithms (such as JPEG, MPEG, and H.264) and discrete wavelet transform (DWT) - based JPEG2000 are discussed with more details. This is followed by discussion on key properties of various transform spaces with special emphasis on the block DCT space and the DWT.

Different types of image and video processing operations performed in the compressed domain are discussed in subsequent chapters. This includes filtering, enhancement and color restoration, image and video resizing, transcoding, etc. In addition, different other applications in the compressed domain such as video and image editing, digital watermarking and steganography, image and video indexing, face detection and identification, etc., are briefly covered in the last chapter.

This book is meant for readers who have gone through a first-level course on digital image processing and are familiar with the basic concepts and tools of image and video processing. However, at the introductory level, the details of compression techniques and properties of the transform domain are not always extensively covered. The first two chapters of this book discuss these issues at considerable depth. For the sake of completeness, an effort has also been made to develop concepts related to compressed domain processing from the very basic level so that a first-time reader also does not have any problem in understanding them.

The book has seven chapters. In the first chapter, the motivation and background for processing images and videos in the compressed domain are discussed. There is also a brief introduction to different popular image and video compression algorithms, notably JPEG, JPEG2000, MPEG-2, MPEG-4, and H.264 standards for lossy image and video compression schemes. Issues related to compressed domain analysis and performance metrics for comparing different algorithms are also elaborated in this chapter.

The second chapter elucidates the definitions and properties of different image transforms, in particular, the discrete fourier transform (DFT), DCT, integer cosine transform (ICT), and DWT. The last three transforms are cho-

sen because of their use in compression technologies. In subsequent chapters, some of the core operations such as filtering, resizing, etc., which find use in various approaches in image and video analysis exploiting these properties, are discussed.

The third chapter considers image filtering in the block DCT domain. In this chapter the convolution and multiplication properties of DCTs are elaborated, followed by discussion on different approaches for computing the filtered response directly in the compressed domain. Typical applications of filtering are also briefly presented in this chapter.

In chapter four, with a general introduction to color processing in the compressed domain, a few representative problems are considered. The first one is related to image enhancement and restoration through saturation and desaturation of colors. In the second case, the problem of color constancy is introduced, and various approaches for solving this problem in spatial and compressed domain are illustrated. A comparative study of different representative schemes is also presented here. Next, the problem of enhancing colors in the block DCT domain is taken into account, and different algorithms for performing this task are discussed.

Chapter five focusses on the image resizing problem in the block DCT space. Different approaches are discussed in this regard. Initially, various techniques for image halving and doubling are discussed. Later, the problem of arbitrary resizing is considered. The chapter also introduces the concept of hybrid resizing and discusses its solution in the block DCT space. The problem of video resizing, in particular video downsampling, is discussed in the next chapter on transcoding.

In chapter six, transcoding of images and videos is discussed. As most of the image and video standards use the DCT, ICT, or DWT for their representation, first, techniques for intertransform conversion are discussed. These are followed by a discussion on various types of transcoding operations. The topics of interest include transcoding of a JPEG2000 image into JPEG, an H.264 video into MPEG-2, and vice versa. The discussion is facilitated with the introduction of different measures related to the performance of a transcoder. The chapter also discusses techniques for altering temporal and spatial resolution of videos by skipping frames at regular intervals and reducing their frame sizes, respectively. At the end of the chapter, error-resilient transcoding of the video stream is also discussed.

There are various other applications of processing of images and videos in the compressed domain. They include different video editing operations such as key frame extraction, caption localization, object recognition, etc. There are also different methods of indexing videos using features computed from the block DCT and DWT spaces. Image and video steganography and watermarking in the compressed domain are also major topics of research in the area of multimedia security. All these different facets of compressed domain analysis are put together in the concluding seventh chapter.

Even after going through several revisions of the text, I always found scopes

for improvements at every iteration. Finally, I had to settle for this version to meet the deadline and other commitments. I would greatly appreciate if the readers of this book, after encountering errors in the printed text, would bring them to my notice.

While working in this area I have been fortunate to have had the guidance and friendship of Professor Sanjit K. Mitra of the University of Southern California, Los Angeles. I take this opportunity to express my deepest gratitude and respect for his constant encouragement and enlightenment. I am also thankful to my colleagues Professor P.K. Biswas and Professor Rajeev Kumar of IIT, Kharagpur, who have worked with me in this area of research. My gratitude also goes to my former students Dr. K. Viswanath, Dr. V. Patil, Mr. Sudhir Porwal, and Ms. T. Kalyani, who contributed in this area at different stages and enriched my understanding of this topic. Professor Shamik Sural of IIT, Kharagpur, went through several versions of this book and greatly helped in improvising the present edition. Without his help and constant encouragement, it would not have been possible for me to complete this book. I also thank Dr. Sreepat Jain of CRC press who invited me to write this book and initiated the project. Later, I received able support from Ms. Aastha Sharma and Ms. Jessica Vakili toward its completion. I am grateful to Jim McGovern who did an extensive proof reading of the manuscript and helped me to correct many typos and grammars in this book. Though I am pretty sure my wife Jhuma and my son Rudrabha will be least interested in reading the book's content, they would be at least happy to see the end of their nightmare due to my late-night intellectual exercises to meet the submission-deadline. I especially thank them for their support, patience, and understanding. Finally, I dedicate this book to my parents, from whom I learned my first lessons with the greatest joy ever I had in my life.

Jayanta Mukhopadhyay
24th September, 2010
IIT Kharagpur, India

MATLAB® is a trademark of The MathWorks, Inc. For product information, please contact:

The MathWorks, Inc. 3 Apple Hill Drive Natick, MA 01760-2098 USA
Tel: 508 647 7000 Fax: 508-647-7001 E-mail: info@mathworks.com
Web: www.mathworks.com

Symbol Description

Notations are defined here in 1-D. The same notations are also used in their extended definitions in 2-D.

Sets and spaces

| | |
|-------------------|---|
| \mathbb{R} | The set of all real numbers. |
| \mathbb{Z} | The set of all integers. |
| \mathbb{N} | The set of all non-negative integers. |
| \mathbb{Z}_N | $\{0, 1, 2, \dots, N - 1\}$ |
| \mathbb{C} | The set of all complex numbers. |
| $L^2(\mathbb{R})$ | The space of all square integrable functions. |
| $L^2(\mathbb{Z})$ | The space of all square integrable functions over the integer grid. |
| $[a, b]$ | $\{x x \in \mathbb{R} \text{ and } a \leq x \leq b\}$ |

Functions, sequences, operators and symbols

| | |
|-------------------------|---|
| $\langle h, g \rangle$ | The inner product of two functions $h(x)$ and $g(x)$. |
| $\vec{h} \cdot \vec{g}$ | The dot product of two vectors \vec{h} and \vec{g} . |
| $a + jb$ | A complex number with j as $\sqrt{-1}$. |
| x^* | The complex conjugate of $x \in \mathbb{C}$. |
| $ x $ | The magnitude of $x \in \mathbb{C}$. |
| $\angle x$ | The phase of $x \in \mathbb{C}$. |
| $\langle x \rangle_N$ | $x \bmod N$ for $x \in \mathbb{Z}$. |
| $ x $ | The absolute value of $x \in \mathbb{R}$. |
| $sign(x)$ | $-1, 0$, and 1 depending on the sign of $x \in \mathbb{R}$. |
| $round(x)$ | The nearest integer approximation of $x \in \mathbb{R}$. |
| $\lfloor x \rfloor$ | The nearest integer, which is less than or equal to $x \in \mathbb{R}$. |
| $\lceil x \rceil$ | The nearest integer, which is greater than or equal to $x \in \mathbb{R}$. |
| $f(x)$ | A continuous function ($x \in \mathbb{R}$) in $L^2(\mathbb{R})$. |
| $\delta(x)$ | The Dirac delta function. |
| $f(n)$ | A discrete function ($n \in \mathbb{Z}$) in $L^2(\mathbb{Z})$. |
| $f^+(n)$ | Positive half of $f(n)$ for $n \geq 0$. |
| $f^p(n)$ | Strict positive half of $f(n)$ for $n > 0$. |
| $x_{de}(m)$ | Even down-sampled sequence of $x(n)$. |
| $x_{do}(m)$ | Odd down-sampled sequence of $x(n)$. |
| $x_{ue}(m)$ | Even up-sampled sequence of $x(n)$. |
| $x_{uo}(m)$ | Odd up-sampled sequence of $x(n)$. |
| $w(n)$ | The conjugate reflection of $w(n) \in \mathbb{C}$. |
| $\ \mathbf{x}\ $ | The Euclidean norm of the vector \mathbf{x} . |
| $f \star h(n)$ | Linear convolution of $f(n)$ and $h(n)$ (or $f(n) \star h(n)$). |
| $f \circledast h(n)$ | Circular convolution of $f(n)$ and $h(n)$ (or $f(n) \circledast h(n)$). |

| | |
|-----------------------|---|
| $f \circledcirc h(n)$ | Skew circular convolution of $f(n)$ and $h(n)$ (or $f(n) \circledcirc h(n)$). |
| $f \bowtie h(n)$ | Symmetric convolution of $f(n)$ and $h(n)$ (or $f(n) \bowtie h(n)$). |
| n_m | Number of multiplications. |
| n_a | Number of additions. |

Transforms

| | |
|--|--|
| $\mathbb{F}(f(x))$ | Fourier transform of $f(x)$. |
| $\hat{f}(j\omega)$ | Fourier transform of $f(x)$ or $\mathbb{F}(f(x))$. |
| $ \hat{f}(j\omega) $ | Magnitude spectrum of $f(x)$. |
| $\theta(\omega)$ | Phase spectrum of $f(x)$. |
| $\mathbb{F}(f(n))$ | The DFT of $f(n)$. |
| $\hat{f}(k)$ | The DFT of $f(n)$ or $\mathbb{F}(f(n))$. |
| $\mathbb{F}_{\alpha,\beta}(f(n))$ | The GDFT of $f(n)$ for $\alpha, \beta \in \{0, \frac{1}{2}\}$. |
| $\hat{f}_{\alpha,\beta}(k)$ | The GDFT of $f(n)$ for $\alpha, \beta \in \{0, \frac{1}{2}\}$, or $\mathbb{F}_{\alpha,\beta}(f(n))$. |
| $\vec{\hat{x}}$ | The DFT of \vec{x} . |
| $\hat{f}_{0,\frac{1}{2}}(k)$ | The Odd Time Discrete Fourier Transform (OTDFT) of $f(x)$. |
| $\hat{f}_{\frac{1}{2},0}(k)$ | The Odd Frequency Discrete Fourier Transform (OFDFT) of $f(x)$. |
| $\hat{f}_{\frac{1}{2},\frac{1}{2}}(k)$ | The Odd Frequency Odd Time Discrete Fourier Transform (O^2DFT) of $f(x)$. |
| $H(z)$ | The z-transform of $h(n)$. |
| $C_{ie}(x(n))$ | Type-i even DCT of $x(n)$ for $i \in \{1, 2, 3, 4\}$. |
| $X_{ie}(k)$ | Type-i even DCT of $x(n)$ for $i \in \{I, II, III, IV\}$ (an alternative notation of $C_{ie}(x(n))$). |
| $C_{io}(x(n))$ | Type-i odd DCT of $x(n)$ for $i \in \{1, 2, 3, 4\}$. |
| $X_{io}(k)$ | Type-i odd DCT of $x(n)$ for $i \in \{I, II, III, IV\}$ (an alternative notation of $C_{io}(x(n))$). |
| $S_{ie}(x(n))$ | Type-i even DST of $x(n)$ for $i \in \{1, 2, 3, 4\}$. |
| $X_{ise}(k)$ | Type-i even DST of $x(n)$ for $i \in \{I, II, III, IV\}$ (an alternative notation of $S_{ie}(x(n))$). |
| $S_{io}(x(n))$ | Type-i odd DST of $x(n)$ for $i \in \{1, 2, 3, 4\}$. |
| $X_{iso}(k)$ | Type-i odd DST of $x(n)$ for $i \in \{I, II, III, IV\}$ (an alternative notation of $S_{io}(x(n))$). |
| $DCT(x)$ | Type II even DCT of x . |
| $DST(x)$ | Type II even DST of x . |

Matrices and operators

| | |
|---------------|---|
| X^T | The transpose of matrix X . |
| X^H | The Hermitian transpose of matrix X . |
| X^{-1} | The inverse of matrix X . |
| $A \otimes B$ | Element wise multiplication of A and B . |
| $[f(k, l)]$ | The matrix formed in such a way that its (k, l) th element is $f(k, l)$. |

| | | | |
|------------------------------|--|------------------|---|
| \mathbf{x} | The column vector formed from $x(n)$ such that i th element of the vector is $x(i)$. | S_N^α | N -point Type α even DST matrix, where $\alpha \in \{I, II, III, IV\}$. |
| $\{\mathbf{x}\}_p^q$ | The column vector formed from \mathbf{x} from its p th element to q th one. | C_N | N -point Type II even DCT matrix. |
| $\mathbb{D}(\mathbf{x})$ | The diagonal matrix whose (i, i) th diagonal element is $x(i)$. | S_N | N -point Type II even DST matrix. |
| $\mathbb{D}_m(\mathbf{x})$ | The diagonal matrix whose m th off diagonal elements are formed from \mathbf{x} in the order of appearances while scanning from left to right and top to bottom. | C_8 | 8-point Type II even DCT matrix. |
| Φ_N | The $N \times N$ flipping matrix. | T_8 | 8-point ICT matrix. |
| Ψ_N | The diagonal matrix $\mathbb{D}(\{(-1)^m\}_{m=0}^{N-1})$. | T_4 | 4-point ICT matrix. |
| \mathbf{F} | The DFT matrix. | Hd_m | Hadamard matrix of size $2^m \times 2^m$. |
| $\mathbf{F}_{\alpha, \beta}$ | The GDFT matrix for $\alpha, \beta \in \{0, \frac{1}{2}\}$. | \mathbb{H}_N | Discrete Haar Transform matrix of size $N \times N$. |
| C_N^α | N -point Type α even DCT matrix, where $\alpha \in \{I, II, III, IV\}$. | P_N | $N \times N$ permutation matrix. |
| | | 0_N | $N \times N$ zero or null matrix. |
| | | $0_{M \times N}$ | $M \times N$ zero or null matrix. |
| | | I_N | $N \times N$ identity matrix. |
| | | J_N | $N \times N$ reverse identity matrix. |
| | | $A_{M, N}$ | DCT block composition matrix for merging M adjacent N -point DCT blocks. |
| | | $B_{M, N}$ | DST block composition matrix for merging M adjacent N -point DCT blocks into a DST block. |

List of Figures

| | | |
|------|--|----|
| 1.1 | Schematic representations of (a) compression and (b) decompression schemes. | 5 |
| 1.2 | Basic steps in processing with compressed images. | 7 |
| 1.3 | Processing in the domain of alternative representation. | 8 |
| 1.4 | Steps in baseline sequential JPEG lossy compression scheme. | 11 |
| 1.5 | A typical quantization table. | 12 |
| 1.6 | DC offset. | 13 |
| 1.7 | Zigzag sequence of AC coefficients. | 14 |
| 1.8 | Encoding steps of a macroblock. | 15 |
| 1.9 | Steps of JPEG2000 compression technique. | 16 |
| 1.10 | Discrete Wavelet Transform and its inverse in 1-D. (From [147] with permission from the author.) | 16 |
| 1.11 | Dyadic decomposition of images: (a) single-level and (b) 3-level. | 18 |
| 1.12 | Code structure of JPEG2000. | 20 |
| 1.13 | Block diagram of a typical video encoder. | 21 |
| 1.14 | Video stream data hierarchy. | 22 |
| 1.15 | Video sequence in an MPEG stream. | 23 |
| 1.16 | Structure of a macroblock. | 24 |
| 1.17 | Prediction of pictures in a GOP sequence. | 24 |
| 1.18 | Coding of predicted pictures. | 26 |
| 1.19 | Coding of bidirectional predicted pictures. | 26 |
| 1.20 | Motion estimation for MPEG-2 video encoder. | 27 |
| 1.21 | An overview of an MPEG-4 video encoder. | 29 |
| 1.22 | Structure of H.264/AVC video encoder. | 33 |
| 1.23 | Basic coding structure of H.264/AVC. | 34 |
| 1.24 | Intra_4 × 4 prediction is computed for samples $a - p$ of a block using samples $A - Q$ | 35 |
| 1.25 | Intra_4 × 4 prediction modes. | 36 |
| 1.26 | Partitioning of an MB. | 37 |
| 1.27 | Image resizing from (a) HDTV (of frame-size 1080 × 1920) to (b) NTSC (of frame-size 480 × 640). | 41 |
| 1.28 | Color image enhancement in the block DCT space: (a) Original and (b) Enhanced image. | 42 |

| | | |
|------|--|-----|
| 2.1 | An example of periodic extension: (a) a finite length sequence and (b) its periodic extension. | 62 |
| 2.2 | Different types of symmetric and antisymmetric extensions at the end of a sequence: (a) Original, (b) WS, (c) HS, (d) WA, and (e) HA. | 69 |
| 2.3 | Symmetrically extended even periodic sequences from a finite length sequence of length 4. | 71 |
| 2.4 | Symmetrically extended odd periodic sequences from a finite length sequence of length 4. | 72 |
| 2.5 | The discrete wavelet transform and its inverse using filter banks.(Reproduced from Figure 1.10 of Chapter 1.) | 93 |
| 2.6 | Dual lifting operations. | 96 |
| 2.7 | Inverse lifting operations. | 97 |
| 2.8 | 2-D DWT and IDWT. | 103 |
| 3.1 | (a) Impulse response, (b) right-half response, (c) input sequence, and (d) zero-padded input sequence. | 116 |
| 3.2 | Removal of blocking artifacts of highly compressed JPEG images using Gaussian filtering in the compressed domain:(a) original image, and (b) compressed at quality factor 10 (JPQM = 3.88), and (c) filtered image (JPQM = 10.74). | 133 |
| 3.3 | Image sharpening using Gaussian filtering with $\lambda = 0.9$ and $\sigma = 2.0$ (JPQM = 9.87). | 134 |
| 4.1 | CIE chromaticity diagram | 139 |
| 4.2 | Saturation and desaturation operation in xy chromaticity space. | 141 |
| 4.3 | Saturation–desaturation in nCbCr space. | 143 |
| 4.4 | Maximum saturation in nCbCr: (a) image22 and (b) maximally saturated image. | 160 |
| 4.5 | Chromaticity points in nCbCr: (a) image22 and (b) maximally saturated image. | 161 |
| 4.6 | (a) image22: Desaturation from the maximally saturated image and (b) corresponding chromaticity plot in the nCbCr space. | 162 |
| 4.7 | image22: (a) Maxsat in DCT and (b) SatDesat in DCT. | 163 |
| 4.8 | Example of color correction using COR-DCT: (a) original image, (b) diagonal correction, and (c) chromatic shift. | 164 |
| 4.9 | Plot of $\eta(x)$ | 164 |
| 4.10 | Plot of $\tau(x)$ | 165 |
| 4.11 | Original images: (a) image7 and (b) image20. | 165 |
| 4.12 | Color enhancement of the images. | 166 |
| 4.13 | Color enhancement of ‘image20’ obtained with more than one iteration: (a) 2 and (b) 3. | 167 |
| 5.1 | Four adjacent spatial domain blocks. | 171 |

| | | |
|------|--|-----|
| 5.2 | Image halving using linear distributive and unitary properties of DCT: (a) original image, (b) bilinear decimation in spatial domain, and (c) downsampled image in the transform domain (58.24 dB with respect to the image in (b) with JPQM as 8.52). | 172 |
| 5.3 | Image halving and doubling using convolution-multiplication property [62]: (a) downsampled image (36 dB with respect to the image in Figure 5.2(b) with JPQM as 9.31) and (b) upsampled image with JPQM as 6.29. | 176 |
| 5.4 | Image halving [36]: four 4×4 approximated DCT coefficients of adjacent blocks are merged into an 8×8 DCT block. | 178 |
| 5.5 | Image halving [103]: four 8×8 are composed and the composed block is approximated to an 8×8 DCT block. | 179 |
| 5.6 | Image halving: (a) IHAC (37.43 dB with respect to the image in Figure 5.2(b) with JPQM as 9.14) and (b) IHCA (36 dB with respect to the image in Figure 5.2(b) with JPQM as 9.31). | 179 |
| 5.7 | Image doubling [36]: an 8×8 DCT block is decomposed into four 4×4 DCT blocks, each of which is approximated to an 8×8 DCT block with zero-padding. | 180 |
| 5.8 | Image doubling [106]: an 8×8 DCT block is approximated to a 16×16 DCT block with zero-padding, which is subsequently decomposed into four 8×8 DCT blocks. | 181 |
| 5.9 | Image doubling through IDDA (JPQM = 5.26). | 182 |
| 5.10 | Image doubling through IDAD (JPQM = 6.88). | 183 |
| 5.11 | Image resizing by a factor of 3×2 : (a) downsampled image (JPQM = 9.19) and (b) upsampled image (JPQM = 5.86). | 188 |
| 5.12 | Conversion of an HDTV frame (1080×1920) to an NTSC frame (480×640): (a) HDTV, (b) NTSC by UDRA (JPQM = 10.21), (c) NTSC by DURA (JPQM = 8.35), and (d) NTSC by HRA (JPQM = 9.92). | 190 |
| 5.13 | Hybrid resizing: (a) by $\frac{2}{3} \times \frac{4}{5}$, (b) by $\frac{3}{2} \times \frac{5}{4}$ (c) by $\frac{2}{3} \times \frac{3}{2}$, and (d) by $\frac{3}{2} \times \frac{2}{3}$. | 194 |
| 6.1 | Forward and inverse DWT in 1-D. (From [147] with permission from the author.) | 197 |
| 6.2 | JPEG2000 to JPEG transcoding for Lena image using the WBTD technique: (a) rate distortion for transcoded image with respect to the JPEG2000 decoded image (for 0.55 bpp encoded Lena image), (b) equivalent rate of the transcoder versus JPEG2000 compression rate, and (c) equivalent PSNR versus equivalent rate. (From [147] with permission from the author.) | 212 |
| 6.3 | Maximum PSNR values versus compression ratio: (a) Lena, (b) Peppers, (c) Girl, and (d) Baboon. (With permission from [147].) | 213 |

| | | |
|------|--|-----|
| 6.4 | Results for the JPEG2000 encoded Lena image:(a) JPEG2000 encoded image at $\text{bpp} = 0.7062$, $\text{PSNR} = 38.47$ dB, (b) JPEG encoded by WBDT-all maximum $\text{PSNR} = 35.21$ dB (at $\text{bpp} = 0.9$ or higher), (c) JPEG encoded image by WBDT-LL maximum $\text{PSNR} = 33.94$ dB (at $\text{bpp} = 0.666$ or higher), and (d) JPEG encoded image by DAT maximum $\text{PSNR} = 29.95$ dB (at $\text{bpp} = 0.784$ or higher). (With permission from [147].) | 214 |
| 6.5 | Single blockwise inverse motion compensation (SBIMC). . . . | 216 |
| 6.6 | Macroblockwise inverse motion compensation (MBIMC). . . . | 218 |
| 6.7 | An integrated scheme for downscaling and inverse motion compensation. | 219 |
| 6.8 | DCT domain based video downscaling system. | 223 |
| 6.9 | Motion Vector propagation from skipped frame. | 226 |
| 6.10 | Cascading pixel domain transcoder. | 228 |
| 6.11 | Transcoding aided by relevant information from the compressed stream. | 228 |
| 6.12 | Hybrid transcoder from H.264 to MPEG-2 (adapted from [102]). | 229 |
| 6.13 | Transcoder for Intra H.264 to Intra MPEG-2 (adapted from [161]). | 232 |
| 6.14 | Error-resilient transcoding in video transmission (adapted from [145]). | 233 |
| 7.1 | Block diagram for computation of caption localization from an I-Frame. | 238 |
| 7.2 | Block diagram for recognition of faces and palmprints. . . . | 241 |
| 7.3 | Block diagram for face detection. | 242 |
| 7.4 | Block diagram for selection of key frames. | 249 |

List of Tables

| | | |
|------|--|-----|
| 1.1 | Daubechies 9/7 analysis and synthesis filter banks for lossy compression | 17 |
| 1.2 | 5/3 analysis and synthesis filter banks for lossless compression | 17 |
| 2.1 | Fourier transforms of a few useful functions | 56 |
| 2.2 | Properties of discrete Fourier transforms of sequences of length N | 64 |
| 2.3 | Different types of DFTs of a sequence $f(n)$ | 65 |
| 2.4 | Summary of observations of symmetric periodic extension of a sequence of length N | 70 |
| 2.5 | Different types of discrete cosine transforms (DCT) | 73 |
| 2.6 | Different types of discrete sine transforms (DST) | 73 |
| 2.7 | Daubechies 9/7 analysis and synthesis filter banks for lossy compression | 95 |
| 2.8 | 5/3 analysis and synthesis filter banks for lossless compression | 95 |
| 2.9 | Complexity comparison of various algorithms for 8×8 DCT | 102 |
| 3.1 | Convolution-multiplication properties for symmetrically extended sequences involving <i>type-II even DCT</i> and <i>type-II even DST</i> | 113 |
| 3.2 | Number of zeroes required to be added at both ends for every type of symmetric extensions | 116 |
| 3.3 | Computational steps and associated complexities of filtering an input DCT block of size N with symmetric or nonsymmetric responses | 125 |
| 3.4 | Per sample computational cost of filtering in 1-D | 126 |
| 3.5 | Per-pixel computational cost of filtering in 2-D | 127 |
| 3.6 | Per-pixel computational cost of filtering in 2-D using the ASC | 128 |
| 3.7 | Per-pixel computational cost of filtering in 2-D using the SNC | 128 |
| 3.8 | Per-pixel computational cost of filtering through convolution in the spatial domain | 129 |
| 3.9 | A noncausal symmetric FIR | 130 |
| 3.10 | A noncausal antisymmetric FIR | 130 |
| 3.11 | A noncausal FIR | 131 |
| 3.12 | A causal FIR | 131 |
| 3.13 | PSNR (in DB) and JPQM values of filtered output using proposed techniques | 131 |

| | | |
|-----|---|-----|
| 4.1 | Performance of different algorithms on the image <i>Bridge</i> . . . | 145 |
| 4.2 | Per-pixel computational cost | 146 |
| 4.3 | List of algorithms for estimating the color components of an illuminant | 149 |
| 4.4 | List of illuminants. | 150 |
| 4.5 | Performance of different techniques of estimating illuminants | 150 |
| 4.6 | Complexities of different algorithms given n_l number of illuminants, n_c as the size of the 2-D chromaticity space, and n number of image pixels | 152 |
| 4.7 | Average performance measures obtained by different color enhancement techniques | 158 |
| 4.8 | Performance measures after iterative application of the CES algorithm on the image 'image20' | 158 |
| 5.1 | (PSNR in DB, JPQM) values after halving and doubling a grey-level image | 183 |
| 5.2 | Per-pixel (of the original image) computational cost of LMDS | 186 |
| 5.3 | Per-pixel computational cost (of the upsampled image) of LMUS | 187 |
| 5.4 | JPQM values of resized images of <i>Watch</i> | 192 |
| 5.5 | Per-pixel computational cost of arbitrary resizing algorithms | 193 |
| 6.1 | Daubechies 9/7 analysis and synthesis filter banks for lossy compression | 201 |
| 6.2 | Per-pixel computational cost of different transcoding approaches. | 210 |
| 6.3 | Distinguishing features of main profiles of MPEG-2 and H.264/AVC | 227 |

Contents

| | | |
|----------|--|----------|
| 1 | Image and Video Compression: An Overview | 1 |
| 1.1 | Compression: Generic Approaches | 2 |
| 1.1.1 | Alternative Representation | 3 |
| 1.1.2 | Quantization | 5 |
| 1.1.3 | Entropy Coding | 5 |
| 1.1.4 | Rate-Distortion Control | 6 |
| 1.2 | Motivation for Processing in the Compressed Domain | 6 |
| 1.3 | Overview of Different Image and Video Compression Techniques and Standards | 8 |
| 1.4 | Image Compression Techniques | 9 |
| 1.4.1 | Baseline Sequential JPEG Lossy Encoding Scheme | 10 |
| 1.4.1.1 | Level Shifting | 10 |
| 1.4.1.2 | Transformation | 10 |
| 1.4.1.3 | Quantization | 11 |
| 1.4.1.4 | Encoding DC Coefficients | 12 |
| 1.4.1.5 | Encoding AC Coefficients | 13 |
| 1.4.1.6 | Entropy Encoding | 13 |
| 1.4.1.7 | Encoding Colors | 14 |
| 1.4.2 | JPEG2000 | 15 |
| 1.4.2.1 | Discrete Wavelet Transform (DWT) | 16 |
| 1.4.2.2 | Quantization | 18 |
| 1.4.2.3 | Bit-Stream Layering, Packetization, and Entropy Coding | 19 |
| 1.4.2.4 | Color Encoding | 19 |
| 1.5 | Video Compression Techniques | 20 |
| 1.5.1 | MPEG-2 | 22 |
| 1.5.1.1 | Encoding Structure | 22 |
| 1.5.1.2 | Frame Types | 23 |
| 1.5.1.3 | Method of Encoding Pictures | 25 |
| 1.5.1.4 | Motion Estimation | 27 |
| 1.5.1.5 | Handling Interlaced Video | 28 |
| 1.5.2 | MPEG-4 | 29 |
| 1.5.2.1 | Video Object Layer | 29 |
| 1.5.2.2 | Background Encoding | 31 |
| 1.5.2.3 | Wavelet Encoding of Still Images | 31 |
| 1.5.3 | H.264/AVC | 31 |