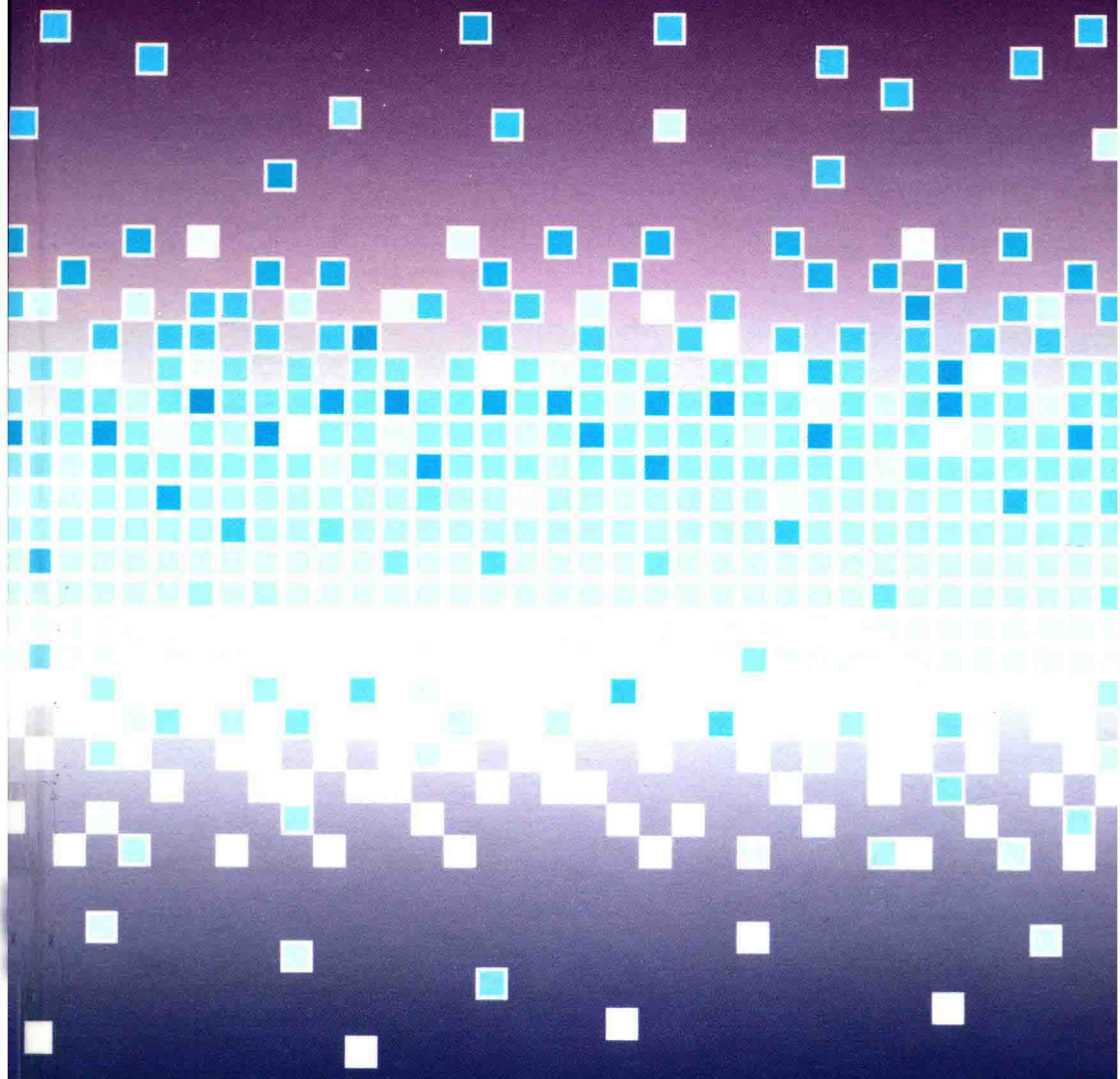
# Image and Video Processing in the Compressed Domain



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

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



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)

# 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

# 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. |
| $\mathbb{Z}_N$ | $\{0, 1, 2, \ldots, N-1\}$ |
| $\mathbb{C}$ | 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. |
| $[a, b]$ | $\{x \mid x \in \mathbb{R} \text{ and } a \leq x \leq b\}$ |

## *Functions, sequences, operators and symbols*

| | |
|---|---|
| $< h, g >$ | 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 $x \in \mathbb{C}$. |
| $< x >_N$ | $x$ **mod** $N$ 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 $x(n)$. |
| $\overline{w(n)}$ | The conjugate reflection of $w(n) \in \mathbb{C}$. |
| $\|\|\mathbf{x}\|\|$ | The Euclidean norm of the vector $\mathbf{x}$. |
| $f \bigstar h(n)$ | Linear convolution of $f(n)$ and $h(n)$( or $f(n) \bigstar h(n)$). |
| $f \circledast h(n)$ | Circular convolution of $f(n)$ and $h(n)$( or $f(n) \circledast h(n)$). |
| $f \textcircled{S} h(n)$ | Skew circular convolution of $f(n)$ and $h(n)$( or $f(n) \textcircled{S} h(n)$). |
| $f \maltese h(n)$ | Symmetric convolution of $f(n)$ and $h(n)$( or $f(n) \maltese h(n)$). |
| $n_m$ | 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 $\mathbb{F}_{\alpha,\beta}(f(n))$. |
| $\hat{\vec{x}}$ | 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 $X$. |
| $A \bigotimes B$ | 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 $x(i)$. |
| $\{\mathbf{x}\}_p^q$ | The column vector formed from $\mathbf{x}$ from its $p$th element to $q$th one. |
| $\mathbb{D}(\mathbf{x})$ | The diagonal matrix whose $(i,i)$th diagonal element is $x(i)$. |
| $\mathbb{D}_m(\mathbf{x})$ | 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. |
| $\Phi_N$ | The $N \times N$ flipping matrix. |
| $\Psi_N$ | The diagonal matrix $\mathbb{D}(\{(-1)^m\}_{m=0}^{N-1})$. |
| $\mathbf{F}$ | The DFT matrix. |
| $\mathbf{F}_{\alpha,\beta}$ | The GDFT matrix for $\alpha, \beta \in \{0, \frac{1}{2}\}$. |
| $C_N^\alpha$ | $N$-point Type $\alpha$ even DCT matrix, where $\alpha \in \{I, II, III, IV\}$. |
| $S_N^\alpha$ | $N$-point Type $\alpha$ even DST matrix, where $\alpha \in \{I, II, III, IV\}$. |
| $C_N$ | $N$-point Type II even DCT matrix. |
| $S_N$ | $N$-point Type II even DST matrix. |
| $C_8$ | 8-point Type II even DCT matrix. |
| $T_8$ | 8-point ICT matrix. |
| $T_4$ | 4-point ICT matrix. |
| $Hd_m$ | Hadamard matrix of size $2^m \times 2^m$. |
| $\mathbb{H}_N$ | Discrete Haar Transform matrix of size $N \times N$. |
| $P_N$ | $N \times N$ 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

# *List of Tables*

# *Contents*