



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ELETRÔNICA E SISTEMAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

VÍTOR DE ANDRADE COUTINHO

**APROXIMAÇÕES DE BAIXA COMPLEXIDADE PARA
TRANSFORMADAS DISCRETAS MULTIDIMENSIONAIS**

Recife

2018

VÍTOR DE ANDRADE COUTINHO

**APROXIMAÇÕES DE BAIXA COMPLEXIDADE PARA
TRANSFORMADAS DISCRETAS MULTIDIMENSIONAIS**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Engenharia Elétrica.

Área de concentração: Comunicações.

Orientador: Prof^o. Dr. Renato José de Sobral Cintra.

Recife

2018

Catálogo na fonte
Bibliotecária Margareth Malta, CRB-4 / 1198

- C871a Coutinho, Vítor de Andrade.
Aproximações de baixa complexidade para transformadas discretas multidimensionais / Vítor de Andrade Coutinho. – 2018.
140 folhas, fig., tabs., abrev. simb.
- Orientador: Prof. Dr. Renato José de Sobral Cintra.
- Tese (Doutorado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia Elétrica, 2018.
Inclui Referências e Apêndices.
1. Engenharia Elétrica. 2. Telecomunicações. 3. Transformadas aproximadas. 4. Transformadas discretas. 5. Transformadas multidimensionais. 6. Compressão de imagem. I. Cintra, Renato José de Sobral. (Orientador). II. Título.
- UFPE
- 621.3 CDD (22. ed.) BCTG/2019-137

VÍTOR DE ANDRADE COUTINHO

**APROXIMAÇÕES DE BAIXA COMPLEXIDADE PARA
TRANSFORMADAS DISCRETAS MULTIDIMENSIONAIS**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Engenharia Elétrica.

Aprovada em: 19/10/2018

BANCA EXAMINADORA

Profº. Dr. Ricardo Menezes Campello de Souza (Examinador Interno)
Universidade Federal de Pernambuco

Profº. Dr. Juliano Bandeira Lima (Examinador Interno)
Universidade Federal de Pernambuco

Profº. Dr. Hélio Magalhães de Oliveira (Examinador Externo)
Universidade Federal de Pernambuco

Profº. Dr. André Leite Wanderley (Examinador Externo)
Universidade Federal de Pernambuco

Profº. Dr. Raydonal Ospina Martínez (Examinador Externo)
Universidade Federal de Pernambuco

Dedico a Glauce, a Henrique, a Clara, a Gusta, a Ana, a Memena e a Tonho.

AGRADECIMENTOS

À minha mãe, Glauce; ao meu pai, Henrique; à minha irmã, Clara; aos meus avós Augusta, Ana Maria, Francisco e Antônio Carlos (*in memoriam*); a Fabíola; a Silvana; pela dedicação, suporte, incentivo, amizade e amor irrestrito.

Ao meu orientador, Dr. Renato J. Cintra, pelas muitas horas de dedicação, disponibilidade, ensinamentos, conselhos e incansável empenho em contribuir com a minha formação e meu futuro profissional.

Ao Dr. Fábio M. Bayer, pelas imprescindíveis contribuições para este trabalho e para a minha formação. Embora — por questões burocráticas — Prof. Bayer não conste como coorientador desta tese de doutorado, desempenhou tal papel *de facto* com dedicação.

Ao Dr. Arjuna Madanayake, pela orientação durante período de estágio de doutorado-sanduíche, pelas contribuições acadêmicas, pela experiência, aprendizado e oportunidades que me proporcionou e contribuição com este trabalho.

A toda a minha família e amigos que me acompanharam durante esta jornada.

À Universidade Federal de Pernambuco, que, ao longo de onze anos, me proporcionou uma formação acadêmica e profissional de excelente qualidade.

A todos os meus professores e colegas da pós-graduação, da graduação e do período escolar, que fizeram parte da minha formação.

Ao CNPq e à CAPES, pelo suporte financeiro.

RESUMO

A presente tese de doutorado aborda aproximações de baixa complexidade computacional para transformadas discretas multidimensionais com aplicações em diferentes contextos. Por proporcionarem algoritmos livres de multiplicação, aproximações para transformadas discretas vêm sendo objeto de pesquisa nos últimos anos. Neste trabalho, é objetivado propor: aproximações para a 2D DTT baseadas na abordagem *pruning* (aproximações podadas), em que coeficientes de alta frequência são descartados de modo a reduzir o custo computacional, com aplicações em padrões de imagem e vídeo; aproximações multidimensionais para a DCT baseadas no formalismo de tensores com aplicações em codificação tridimensional de vídeo e rastreamento visual de objetos; e aproximações para a DHT e 3D DHT visando codificação de imagens médicas. Os métodos aproximados podados propostos para a 2D DTT mostraram reduções de complexidade entre 50% e 80% em relação ao método usual, com uma perda de desempenho em compressão tipo JPEG na faixa de 3% a 6%, em média. As versões aproximadas para a 3D DCT apresentaram 100% de redução de complexidade multiplicativa e entre 17% e 52% de redução de complexidade aditiva a uma perda de desempenho em compressão de vídeo 3D menor que 7%. Por sua vez, as aproximações propostas para a 3D DHT apresentaram 100% de redução de complexidade multiplicativa, embora custos aditivos iguais ou superiores ao dos métodos exatos, com redução de desempenho em compressão de imagens médicas menor que 2%. De forma geral, as simulações de compressão de imagem e vídeo e a análise de complexidade computacional sugerem que os métodos propostos apresentam um *trade-off* favorável entre custo computacional e desempenho de compressão, proporcionando um leque de ferramentas que atende à crescente demanda por sistemas com restritos custo computacional e autonomia energética.

Palavras-chave: Engenharia elétrica. Telecomunicações. Transformadas aproximadas. Transformadas discretas. Transformadas multidimensionais. Compressão de imagem.

ABSTRACT

The current doctoral thesis addresses low-complexity multidimensional discrete transform approximations with applications in different contexts. Because they can provide multiplierless algorithms and multiplication, discrete transform approximations have been subject of research in recent years. In the current work, we aim at proposing: DTT approximations based on pruning schemes, which consist in discarding high-frequency coefficients aiming at reducing the overall computational cost, with applications in image and video coding; multidimensional DCT approximations based on tensor formalism with applications in 3D-based video coding and visual tracking; and DHT and 3D DHT approximations for medical image compression. The proposed pruned 2D DTT approximations presented computational complexity reduction from 50% to 80% compared to the usual method, with performance decrease in JPEG-like scheme from 3% to 6%, in average. The proposed 3D DCT approximations showed multiplicative and additive complexity reduction of 100% and additive complexity reduction from 17% to 52%, at a performance loss lower than 7%. Furthermore, the proposed 3D DHT approximations presented 100% of multiplicative complexity reduction, although equal or higher additive costs, with performance decrease smaller than 2%. Generally, the image and video compression simulations and the computational complexity analysis suggest the proposed methods present a favorable trade-off between performance and computational cost, being suitable to attend the increasing demand for systems with strict computational cost and power consumption constraints.

Keywords: Electrical Engineering. Telecommunications. Approximate transforms. Discrete transforms. Multidimensional transforms. Image compression.

LISTA DE FIGURAS

Figura 1 – Algoritmos rápidos para os aproximações podadas para a DTT propostas para cada parâmetro de poda K	47
Figura 2 – Algoritmos rápidos para os aproximações podadas para a DTT propostas para cada parâmetro de poda K mostrando as linhas de fluxos podadas, em vermelho.	48
Figura 3 – Métricas de desempenho em simulação tipo JPEG para a 2D DTT exata, a 2D DTT aproximada e as 2D DTT aproximadas podadas para cada fator de poda K	50
Figura 4 – Métricas de desempenho em compressão de imagem tipo JPEG para os métodos podados propostos para a 2D DTT aproximada e métodos podados para a DCT exata para $QF = 25$	53
Figura 5 – Avaliação qualitativa de desempenho para compressão de imagem ‘lena’ com $QF = 25$ para a 2D DCT e 2D DTT exatas, 2D DTT aproximada e suas versões podadas para cada parâmetro K	54
Figura 6 – Avaliação qualitativa de desempenho para compressão de imagem ‘mandrill’ com $QF = 25$ para a 2D DCT e 2D DTT exatas, 2D DTT aproximada e suas versões podadas para cada parâmetro K	55
Figura 7 – Curvas de desempenho para H.264/AVC original e versões modificadas empregando as aproximações podadas para a 2D DTT para diferentes parâmetros de poda K	56
Figura 8 – Curvas de desempenho para H.265/HEVC original e versões modificadas empregando as aproximações podadas para a 2D DTT para diferentes parâmetros de poda K	57
Figura 9 – Distribuição de energia por coeficiente para os métodos propostos para a 2D DTT e métodos usuais.	59
Figura 10 – Concentração de energia por zona para cada parâmetro de poda K para os métodos propostos e métodos usuais.	60

Figura 11 – Arquiteturas realizadas para aproximações podadas para a 2D DTT com alguns fatores de poda K	62
Figura 12 – Transformadas 3D ótimas para diferentes parâmetros da função custo. . . .	80
Figura 13 – Medidas de desempenho em compressão de vídeo para a 3D DCT exata e diversas aproximações 3D derivadas.	86
Figura 14 – Comparação qualitativa para o primeiro quadro da sequência de vídeo “foreman” empregando a 3D DCT exata e diversas aproximações 3D propostas.	87
Figura 15 – Avaliação qualitativa para rastreamento da sequência de vídeo “animal” . .	90
Figura 16 – Métrica PBM para cada r -ésimo quadro da sequência de vídeo “animal” . .	91
Figura 17 – Algoritmo rápido para a DHT de comprimento oito.	101
Figura 18 – PSNR e SSIM médios versus <i>bitrate</i> para a 3D DHT exata e cada método aproximado proposto em compressão de arranjos XA.	107
Figura 19 – PSNR e SSIM médios versus <i>bitrate</i> para a 3D DHT exata e cada método aproximado proposto em compressão de arranjos MR.	108
Figura 20 – PSNR e SSIM relativos versus <i>bitrate</i> dos métodos propostos comparados com a 3D DHT exata em compressão de arranjos XA.	108
Figura 21 – PSNR e SSIM relativos versus <i>bitrate</i> dos métodos propostos comparados com a 3D DHT exata em compressão de arranjos MR.	109
Figura 22 – Avaliação qualitativa para um slice de MR comprimido a 4 bpv empregando a 3D DHT exata e várias versões aproximadas propostas.	111
Figura 23 – Avaliação qualitativa para um slice de XA comprimido a 4 bpv empregando a 3D DHT exata e várias versões aproximadas propostas.	112

LISTA DE TABELAS

Tabela 1 – Desvios de diagonalidade associados às inversas generalizadas e às matrizes transpostas e a diferença percentual para cada 2D DTT aproximada podada com parâmetro de poda K	45
Tabela 2 – Complexidade das aproximações original e podadas para a 2D DTT de dimensões 8×8 e redução percentual relativa à 2D DTT exata	49
Tabela 3 – Consumo de recursos de <i>hardware</i> e potência para os métodos propostos para diferentes fatores de poda K	63
Tabela 4 – Reduções percentuais de AT e AT^2 de cada 2D DTT aproximada podada com fator de poda K em comparação com a DTT aproximada não podada ($K = 8$)	63
Tabela 5 – Consumo de recursos de <i>hardware</i> e potência para a DCT exata podada para diferentes fatores de poda K	63
Tabela 6 – Reduções percentuais de AT e AT^2 de cada aproximação para a 2D DTT com fator de poda K em comparação com a DCT exata podada também com fator de poda K	64
Tabela 7 – Aproximações para a DCT de comprimento $N = 8$ encontradas na literatura	73
Tabela 8 – Eficiência de codificação e complexidade computacional para a 3D DCT exata diversas aproximações para a 3D DCT derivadas	76
Tabela 9 – Reduções percentuais de eficiência de codificação e complexidade computacional comparado com a 3D DCT exata (levando em conta o algoritmo 3D VR DIF)	81
Tabela 10 – Nível de movimento das sequências de vídeos consideradas	88
Tabela 11 – Desempenho dos métodos propostos para a 3D DCT para cada sequência de vídeo e correlação com nível de movimento	89
Tabela 12 – Representação e complexidade dos parâmetros β_m para cada matriz aproximada para a DHT	102

Tabela 13 – Métricas e complexidades computacionais para as matrizes aproximadas obtidas para a DHT	102
Tabela 14 – Complexidade computacional para os algoritmos para a 3D DHT e versões aproximadas de dimensões $8 \times 8 \times 8$	107

LISTA DE ABREVIATURAS E SIGLAS

AVC	padrão de codificação <i>advanced video coding</i> (H.264)
BAS	série de DCTs aproximadas de Bouguezel-Ahmad-Swamy
CIF	Formato padronizado <i>Common Intermediate Format</i>
CSD	representação <i>canonical-signed-digit</i>
DCT	transformada discreta do cosseno
DFT	transformada discreta do Fourier
DHT	transformada discreta de Hartley
DTT	transformada discreta de Tchebichef
FPGA	arranjo de portas programáveis em campo ou <i>field-programmable gate array</i>
HEVC	padrão de codificação de vídeo <i>High Efficiency Video Coding</i>
HVS	sistema visual humano ou <i>human visual system</i>
IADCT	<i>improved approximate DCT</i>
ICT	DCT inteira
JPEG	padrão de codificação de imagem estática <i>Joint Photographic Experts Group</i>
KLT	transformada de Karhunen-Loève
LODCT	DCT aproximada de Lengwehasatit-Ortega
<i>m</i> D DCT	DCT multidimensional ou <i>m</i> -dimensional
<i>m</i> D DHT	DHT multidimensional ou <i>m</i> -dimensional
<i>m</i> D DTT	DTT multidimensional ou <i>m</i> -dimensional

MPEG	padrão de codificação de vídeo <i>Moving Picture Experts Group</i>
MR	ressonância magnética
MRDCT	DCT arredondada modificada ou <i>modified round-off</i> DCT
PCA	análise de componentes principais
RDCT	DCT arredondada ou <i>round-off</i> DCT
SDCT	DCT sinalizada ou <i>signed</i> DCT
SDHT	DHT especial ou <i>special</i> DHT
VR DIF	algoritmo dizimação na frequência de base vetorial, ou <i>vector-radix decimation-in-frequency</i> , para a 3D DCT
XA	angiografia por raios-X

LISTA DE SÍMBOLOS

\mathbb{R}	conjunto dos números reais
\mathbb{C}	conjunto dos números complexos
\mathbb{F}	conjunto genérico que pode representar tanto \mathbb{C} quanto \mathbb{R}
\mathbb{Q}	conjunto dos números racionais
\mathbb{Z}	conjunto dos números inteiros
\mathbb{N}	conjunto dos números naturais
\mathbb{P}	conjunto $\{0, \pm 1, \pm 2, \pm \frac{1}{2}\}$
\emptyset	conjunto vazio
\mathbf{I}_N	matriz identidade de ordem N
$\delta_{k\hat{k}}$	delta de Kronecker
\top	transposição de uma matriz
$*$	complexo conjugado (escalar) ou transposição conjugada (matriz)
MSE	Erro médio quadrático
$\text{tr}(\cdot)$	traço de uma matriz
C_g	ganho de codificação ou <i>coding gain</i>
$\text{sum}(\cdot)$	retorna a soma dos elementos do arranjo de seu argumento
\odot	produto matricial de Hadamard
$\partial(\cdot)$	desvio de diagonalidade
$\ \cdot\ $	norma euclidiana de um vetor

$\ \cdot\ _F$	norma matricial de Frobenius
$\text{diag}(\cdot)$	retorna uma matriz diagonal com os elementos da diagonal do argumento
$\text{PSNR}(\cdot)$	relação sinal ruído de pico
$\text{MSE}_s(\cdot)$	erro médio quadrático de sinal
QF	fator de qualidade para quantização em simulação tipo JPEG
QP	parâmetro de quantização para codificadores de vídeo H.264 e HEVC
$\text{SSIM}(\cdot)$	índice de similaridade estrutural
\times_i	produto em modo i -ésimo
$\text{PBM}(\cdot)$	métrica baseada em posição ou <i>position-based measure</i>

SUMÁRIO

1	INTRODUÇÃO	19
1.1	REVISÃO DA LITERATURA E MOTIVAÇÃO	19
1.2	OBJETIVOS	28
1.3	ESTRUTURA DO DOCUMENTO	29
2	APROXIMAÇÕES PODADAS PARA A TRANSFORMADA DISCRETA DE TCHEBICHEF BIDIMENSIONAL	31
2.1	REVISÃO MATEMÁTICA	33
2.1.1	Momentos Baseados em Funções Contínuas e Discretas	33
2.1.2	Transformadas Discreta Multidimensionais	34
2.1.3	Transformada Discreta de Tchebichef	38
2.1.4	Aproximações para Transformadas Discretas e Medidas de Desempenho	39
2.1.5	Uma Aproximação de Baixa Complexidade para a DTT	42
2.2	APROXIMAÇÕES PODADAS PROPOSTAS	43
2.2.1	Definição Matemática	43
2.2.2	Análise de Complexidade	45
2.3	APLICAÇÃO EM CODIFICAÇÃO DE IMAGEM E VÍDEO	49
2.3.1	Compressão de Imagem Tipo JPEG	49
2.3.2	Compressão de Vídeo	53
2.3.2.1	H.264/AVC	56
2.3.2.2	H.265/HEVC	57
2.4	ANÁLISE E DISTRIBUIÇÃO DE ENERGIA POR COEFICIENTE	58
2.5	IMPLEMENTAÇÃO EM FPGA	60
3	APROXIMAÇÕES PARA A DCT MULTIDIMENSIONAL	65
3.1	REVISÃO MATEMÁTICA	66
3.1.1	1D e 2D DCT	66

3.1.2	3D DCT e Tensores de Alta Ordem	67
3.1.3	Aproximações para a DCT	71
3.2	DCT MULTIDIMENSIONAL APROXIMADA	74
3.2.1	Definição Matemática	74
3.2.2	Análise de Complexidade	76
3.2.3	Análise de <i>Trade-off</i>	78
3.3	CODIFICAÇÃO DE VÍDEO ENTRE QUADROS BASEADA EM APROXIMAÇÕES PARA A 3D DCT	80
3.3.1	Etapa de Quantização Modificada para Compressão de Vídeo	81
3.3.2	Simulação de Compressão de Vídeo	84
3.4	SISTEMA DE RASTREAMENTO VISUAL BASEADO EM UMA APROXIMAÇÃO PARA A 3D DCT	89
4	APROXIMAÇÕES PARA A DHT TRIDIMENSIONAL E APLICAÇÕES EM COMPRESSÃO DE IMAGEM MÉDICA	94
4.1	DHT: REVISÃO MATEMÁTICA E MÉTODOS APROXIMADOS PROPOSTOS	96
4.1.1	DHT Unidimensional	96
4.1.2	Aproximações para a DHT de Comprimento Oito Propostas	99
4.2	DHT TRIDIMENSIONAL: REVISÃO, DEFINIÇÃO MATEMÁTICA E FORMALIZAÇÃO PARA APROXIMAÇÕES	101
4.2.1	DHT Tridimensional Exata	102
4.2.2	Formalização Matemática para o Caso Tridimensional	104
4.2.3	Análise de Complexidade	105
4.3	APLICAÇÃO EM COMPRESSÃO DE IMAGENS MÉDICAS	106
5	CONCLUSÕES	113
5.1	CONTRIBUIÇÕES	114
5.2	TRABALHOS FUTUROS	116

REFERÊNCIAS	118
APÊNDICE A – IMAGENS EMPREGADAS NA SIMULAÇÃO DE COM- PRESSÃO JPEG	136
APÊNDICE B – TRABALHOS PRODUZIDOS E PUBLICADOS . . .	138

1 INTRODUÇÃO

Neste capítulo, serão apresentados os objetivos e motivação do presente trabalho, assim como uma revisão da literatura e a estrutura do documento.

1.1 REVISÃO DA LITERATURA E MOTIVAÇÃO

Devido ao uso massivo de tecnologias digitais nas últimas décadas, transformadas discretas se tornaram ferramentas matemáticas importantes na área de processamento digital de sinais (AHMED; RAO, 1975; RAO; YIP, 2014; BRITANAK; YIP; RAO, 2007; RAO; YIP, 2001). A aplicabilidade de transformadas discretas em contextos práticos é garantida pela existência de algoritmos rápidos, que proporcionam um cômputo em intervalos de tempo substancialmente menores do que quando implementadas pela definição matemática (BLAHUT, 2010).

Geralmente, a complexidade computacional de um algoritmo é dominada pela contagem de operações de multiplicação, uma vez que sua implementação usualmente demanda mais tempo de execução, área de circuito e energia (DESCHAMPS; BIOUL; SUTTER, 2006, p. 359). A teoria da complexidade multiplicativa para algoritmos rápidos em processamento digital de sinais foi desenvolvida em (HEIDEMAN, 1988). Em tal trabalho, são derivados limites mínimos teóricos para a complexidade multiplicativa de diversas transformadas discretas, tais como a transformada discreta de Fourier (DFT), a transformada discreta do cosseno (DCT) e a transformada discreta de Hartley (DHT).

Transformadas discretas vêm sendo empregadas em uma variedade de aplicações, tais como processamento de imagem e fala (RAO; HWANG, 1996; BHASKARAN; KONSTANTINIDES, 1997; DING; SOON; YEO, 2011; KAFAI; ESHGHI; BHANU, 2014), reconhecimento de padrões (KIRBY; SIROVICH, 1990; JUEFEI-XU; SAVVIDES, 2014), sistemas de comunicação (HARMUTH, 2012; ZHOU et al., 2016) e filtros digitais (PRATT, 1972; SURESH; SREENIVAS, 2008; BÄCKSTRÖM, 2015). Particularmente, sistemas de compressão de imagem e vídeo representam uma área importante do campo de processamento de imagem (RAO; YIP, 2001). O propósito de compressão de imagem é reduzir o número de *bits* empregados na

representação dos dados da imagem, salvando largura de banda em sistemas de transmissão e memória física necessária para armazenamento, porém preservando uma qualidade de imagem satisfatória (LIM, 1990).

O difundido padrão de codificação de imagem estática JPEG (WALLACE, 1992) emprega a abordagem de transformada de bloco (LIM, 1990, p. 642), que consiste em computar uma transformação linear bidimensional com alta capacidade de decorrelação de dados (BHASKARAN; KONSTANTINIDES, 1997). Neste contexto, a transformada de Karhunen-Loève (KLT) apresenta a propriedade de diagonalizar a matriz de covariância do sinal transformado e proporciona, portanto, decorrelação ótima de dados (BRITANAK; YIP; RAO, 2007). No entanto, a matriz — ou núcleo — de transformação da KLT varia de acordo com propriedades estatísticas dos dados de entrada. Consequentemente, algoritmos rápidos para a KLT não são alternativas adequadas para implementações práticas (RAO; YIP, 2014, p. 31).

A DCT surgiu como uma alternativa para o cômputo da KLT que apresenta núcleo de transformação fixo (AHMED; NATARAJAN; RAO, 1974) e propriedades de compactação de energia e de decorrelação de dados (RAO; YIP, 2001). Matematicamente, a DCT representa uma aproximação assintótica para a KLT quando os dados de entrada podem ser modelados como um processo markoviano de primeira ordem, tais como imagens naturais (BRITANAK; YIP; RAO, 2007). Uma vez que o núcleo de transformação da DCT não depende dos dados de entrada, algoritmos rápidos são factíveis. O algoritmo rápido pioneiro sugerido em (AHMED; NATARAJAN; RAO, 1974) se baseia na relação entre a DCT e a DFT e emprega um algoritmo rápido para a DFT para o cômputo da DCT.

A depender da natureza dos dados, algumas aplicações requerem o cálculo de transformadas multidimensionais (GONZALEZ; WOODS, 2001; CHO, 1991; BOUSSAKTA, 2004; BOZINOVIĆ; KONRAD, 2005; DANCIU et al., 2012; XUE et al., 2015). Algoritmos rápidos para as transformadas multidimensionais usualmente exploram a propriedade de separabilidade de núcleo (GONZALEZ; WOODS, 2001; BOUSSAKTA, 2004; SERVAIS; JAGER, 1997; SONG; LI, 2013) e aplicam sucessivamente algoritmos de sua versão unidimensional para o cômputo da transformada de múltiplas dimensões (BLAHUT, 2010; ZENG; BI; LEYMAN, 2000a). Tal técnica é frequentemente referida como método de *decomposição linha-*

coluna (MADISETTI, 2009; ZENG; BI; LEYMAN, 2000a), pois o algoritmo unidimensional é aplicado calculando sucessivas transformações das linhas e colunas do arranjo de entrada. No contexto de compressão de imagem estática, tal qual o padrão JPEG (WALLACE, 1992), a DCT bidimensional (2D DCT) é aplicada como a transformada de bloco (RAO; YIP, 2001) mencionada anteriormente. Em padrões de codificação de vídeo, tais como MPEG (WALLACE, 1992; International Organisation for Standardisation, 1994), H.261 (International Telecommunication Union, 1990) H.263 (International Telecommunication Union, 1995), H.264 (TEAM, 2003), HEVC (International Telecommunication Union, 2013), a 2D DCT é empregada para descorrelação de cada quadro de vídeo.

Devido à sua vasta gama de aplicações, uma série de algoritmos rápidos para a DCT foram propostos desde sua introdução (CHEN; SMITH; FRALICK, 1977; WANG, 1984; LEE, 1984; VETTERLI; NUSSBAUMER, 1984; SUEHIRO; HATERI, 1986; HOU, 1987b; ARAI; AGUI; NAKAJIMA, 1988; LOEFFLER; LIGTENBERG; MOSCHYTZ, 1989; FEIG; WINOGRAD, 1992), sendo um campo de pesquisa que vem sendo trabalhado intensamente. Por ser esta empregada nos diversos padrões de codificação de imagem e vídeo mencionados anteriormente, um foco especial é dado à DCT de comprimento oito, em particular. De fato, o mínimo teórico para a complexidade multiplicativa estabelecido em (HEIDEMAN, 1988) para a DCT de tal comprimento já foi atingido pelo algoritmo de Loeffler (LOEFFLER; LIGTENBERG; MOSCHYTZ, 1989).

Uma vez que o desenvolvimento de algoritmos rápidos para a DCT de comprimento oito é um campo de pesquisa já bem avançado, propor novos algoritmos para a DCT que possam oferecer melhorias significativas em termos de complexidade computacional torna-se cada vez mais difícil. Neste cenário, diferentes técnicas para reduzir ainda mais o custo computacional da DCT foram consideradas, tais como as DCTs inteiras (ICTs) (FONG; CHAM, 2012; BRITANAK; YIP; RAO, 2007; CHEN; ORAINTARA; NGUYEN, 2000; YOKOTANI et al., 2006; HNATIV, 2014), a técnica binDCT baseada em esquemas *lifting* (LIANG; TRAN, 2001; TRAN, 2000; CHEN et al., 2002), as DCTs aproximadas (HAWHEEL, 2001; LENGWEHASATIT; ORTEGA, 2004; CINTRA; BAYER, 2011; BAYER; CINTRA, 2012; BOUGUEZEL; AHMAD; SWAMY, 2008; BOUGUEZEL; AHMAD; SWAMY, 2009; BOUGUEZEL; AHMAD;

SWAMY, 2013; POTLURI et al., 2014; CINTRA, 2011; CINTRA; BAYER; TABLADA, 2014; MADANAYAKE et al., 2015), os algoritmos podados (*pruning* ou *pruned transforms*) (WANG, 1991; MAKKAOUI; LECUIRE; MOUREAUX, 2010; LECUIRE; MAKKAOUI; MOUREAUX, 2012), assim como tais abordagens combinadas (KOUADRIA et al., 2013; COUTINHO et al., 2016; CINTRA et al., 2016; MECHOUK et al., 2016).

De fato, o padrão de codificação de vídeo HEVC adota uma 2D DCT inteira como um passo chave para descorrelação espacial (International Telecommunication Union, 2013; POURAZAD et al., 2012). Frequentemente, transformadas baseadas em representação inteira, tais como a ICT, permitem implementações livre de multiplicação empregando racionais diádicos (BRITANAK; YIP; RAO, 2007, p. 181) e a representação *canonical-signed-digit* (CSD) (LEE, 1997). Tal representação converte uma multiplicação em diversas adições, o que é vantajoso, dado que operações de multiplicação são consideravelmente mais custosas do que adições. Por outro lado, a complexidade aditiva de tais métodos tende a crescer substancialmente.

Também no contexto de codificação HEVC, aproximações de baixa complexidade aditiva para a 2D DCT e livres de multiplicação foram empregadas com sucesso, atingindo um desempenho competitivo a um custo computacional notavelmente reduzido (POTLURI et al., 2014; COUTINHO et al., 2016; CINTRA et al., 2016). Frequentemente, os vetores de base das transformadas aproximadas consideram elementos do conjunto $\mathbb{P} = \{0, \pm 1, \pm 2, \pm \frac{1}{2}\}$. Uma operação de multiplicação pelos coeficientes de tal conjunto pode ser implementada apenas com deslocamentos de *bits* (*shifts*) e sem incremento na complexidade aditiva.

Genericamente, aproximações para a DCT são transformações lineares que apresentam, por construção, baixo custo computacional, preservando propriedades da DCT *exata*¹, tais como compactação de energia e capacidade de descorrelação. Diferentemente da DCT exata, aproximações não estão sujeitas aos limites teóricos estabelecidos para a complexidade multiplicativa (HEIDEMAN, 1988) e representam um campo de pesquisa em desenvolvimento.

Diversas aproximações para a DCT livres de multiplicação e com baixo custo computacional vêm sendo propostas desde a introdução da pioneira *DCT sinalizada* (SDCT) (HAWHEEL,

¹ Neste trabalho, o termo “transformada exata” é empregada para distinguir uma transformada original de suas versões “aproximadas”.

2001). Alguns exemplos de aproximações para a DCT de comprimento oito relevantes na literatura são: a DCT aproximada de Lengwehasatit-Ortega (LODCT) (LENGWEHASATIT; ORTEGA, 2004), a série de aproximações de Bouguezel-Ahmad-Swamy (BAS) (BOUGUEZEL; AHMAD; SWAMY, 2008; BOUGUEZEL; AHMAD; SWAMY, 2009; BOUGUEZEL; AHMAD; SWAMY, 2010; BOUGUEZEL; AHMAD; SWAMY, 2013), a DCT arredondada (RDCT) (CINTRA; BAYER, 2011), a DCT arredondada modificada (MRDCT) (BAYER; CINTRA, 2012) e a *improved* DCT aproximada (IADCT) (POTLURI et al., 2014).

Por sua vez, os esquemas podados consistem em evitar o cálculo de coeficientes dos vetores de entrada e/ou saída que: (i) possuam baixa concentração de energia ou pouco conteúdo de informação (MARKEL, 1971); e/ou (ii) tenham alta probabilidade de serem descartados numa etapa subsequente (WANG, 1991), tal qual a etapa de quantização em sistemas de codificação de imagem e vídeo (LIM, 1990, p. 591). O primeiro algoritmo podado surgiu em 1971 para a DFT, ainda antes do surgimento da DCT (MARKEL, 1971). Tal algoritmo descarta elementos do vetor de entrada, eliminando as operações aritméticas que envolvem tais coeficientes; sendo aplicável em interpolação, análise e síntese de voz e aproximação por mínimos quadrados com polinômios trigonométricos. Tal técnica foi vastamente explorada para o caso da DFT exata em diferentes contextos (MARKEL, 1971; ROCHE, 1992; FAN; SU, 2007; MEDINA-MELENDEZ; ARIAS-ESTRADA; CASTRO, 2009; BOUGUEZEL; AHMAD; SWAMY, 2003).

Por conta da propriedade da compactação da energia da DCT e pela etapa de quantização de esquemas de codificação de imagem e vídeo descartar coeficientes de alta frequência, algoritmos podados são particularmente proveitosos neste contexto. Algoritmos podados para a DCT são obtidos evitando o cálculo de componentes de alta frequência, que em geral concentram menos energia. Métodos podados foram propostos para o caso da DCT exata (WANG, 1991; MAKKAOUI; LECUIRE; MOUREAUX, 2010; LECUIRE; MAKKAOUI; MOUREAUX, 2012). A combinação de métodos podados e aproximados para o caso da DCT foi estudada pelo autor deste trabalho e colaboradores em trabalhos prévios (COUTINHO et al., 2016; CINTRA et al., 2016; COUTINHO et al., 2015).

Embora seja largamente empregada em esquemas de compressão de imagem e vídeo,

a DCT não é a única alternativa para métodos baseados em transformada de bloco. Particularmente, esquemas de compressão tipo JPEG baseados na transformada discreta de Tchebichef (DTT) provaram ser alternativas realísticas (ISHWAR; MEHER; SWAMY, 2008; PRATTIPATI et al., 2013; PRATTIPATI; SWAMY; MEHER, 2013; ABU et al., 2010; OLIVEIRA et al., 2015; OLIVEIRA et al., 2016). A DTT é baseada no cômputo dos momentos de Tchebichef², que são derivados a partir dos polinômios discretos de Tchebichef e que são capazes de redução de redundância em imagens naturais (COUTINHO; CINTRA; BAYER, 2017). Tais momentos foram introduzidos em (MUKUNDAN; ONG; LEE, 2001), em que se sugere que o uso de polinômios ortogonais discretos ao invés dos usuais polinômios contínuos — tais como os momentos de Legendre — pode melhorar o desempenho em reconstrução de imagem uma vez que a implementação de momentos discretos não requer uma aproximação numérica para integrais contínuas. Uma comparação entre a reconstrução de imagem empregando diferentes bases ortogonais como núcleo de transformadas foi realizada em (HUNT; MUKUNDAN, 2004), onde foi mostrado que a DCT e a DTT apresentam desempenho similar para reconstrução de imagens com as mesmas propriedades.

Em (ISHWAR; MEHER; SWAMY, 2008), foi derivado um algoritmo rápido para a DTT de dimensão 4×4 e aplicado em um esquema de compressão de imagem. O algoritmo em questão se baseia em uma representação inteira e não requer operações de multiplicação a um custo de complexidade aditiva incrementado. Em (PRATTIPATI et al., 2013), foi proposto um algoritmo rápido para a DTT de dimensão 8×8 baseado em uma representação inteira. Além disso, o experimento computacional detalhado em tal trabalho mostra que a DTT apresenta desempenho de compressão de imagem comparável com a ICT.

Embora os trabalhos mencionados anteriormente apresentem algoritmos livres de multiplicação para o cômputo da DTT exata, estes ainda requerem um número elevado de adições devido ao uso da representação CSD (HARTLEY, 1991). Sendo assim, esquemas que reduzam substancialmente o custo aditivo do cômputo da DTT vêm sendo perseguidos.

Em (KOUADRIA et al., 2017), foi proposto um algoritmo podado para a DTT exata.

² Nome original em russo. Também são encontradas transliterações tais como “Chebyshev”, “Tchebychev”, “Tchebycheff”, entre outras. Para manter compatibilidade com a literatura relacionada, opta-se nesse trabalho por “Tchebichef”.

Aproximações de baixa complexidade para a DTT foram propostas em (OLIVEIRA et al., 2015; OLIVEIRA et al., 2016) para aplicações de codificação de imagem de baixa potência. Tais aproximações apresentam algoritmos rápidos livres de multiplicação e com mais baixa complexidade aditiva do que os algoritmos para a DTT exata. Em particular, a aproximação proposta em (OLIVEIRA et al., 2016) apresenta a menor complexidade da literatura, ao nosso melhor conhecimento. Embora ambas as abordagens de transformadas aproximadas e algoritmos podados venham sendo vastamente estudadas (LECUIRE; MAKKAOU; MOUREAUX, 2012; KOUADRIA et al., 2013; COUTINHO et al., 2016), a aplicação de tais técnicas de forma combinada para o caso da DTT é um campo inexplorado, ao nosso melhor conhecimento.

Por sua vez, a DHT foi introduzida como uma alternativa para se computar eficientemente a DFT (BRACEWELL, 1983). A DHT apresenta as vantagens de apresentar núcleo de transformação real e de ser uma involução, isto é, as transformadas diretas são idênticas (exceto por um fator de escalonamento) (BRACEWELL, 1983; GONZALEZ; WOODS, 2001).

Além de ser empregada como uma alternativa para a DFT, a DHT inerentemente também apresenta um leque de aplicações práticas, tais como processamento de imagem (PAIK; FOX, 1988), compressão de imagem (VILLASENOR, 1993; SUNDER; ESWARAN; SRI-RAAM, 2006), criptografia de imagem (LIU et al., 2010), interpolação de sinal (AGBINYA, 1987; MAHARANA; MEHER, 1997; TSENG; LEE, 2009), esquemas de modulação multi-tom (OUYANG et al., 2012; WANG; CHANG, 2001) e marca d'água digital (MANDAL; GHOSAL, 2013). Um conjunto de algoritmos rápidos e arquiteturas de *hardware* está disponível na literatura (MEHER; SRIKANTHAN; PATRA, 2006; BRACEWELL, 1984; SORENSEN et al., 1985; GRIGORYAN, 2004; GUO, 2000; HOU, 1987a).

Apesar do seu amplo emprego em padrões de compressão de vídeos, transformadas bidimensionais não levam em consideração a correlação entre quadros consecutivos. Em geral, padrões de vídeos lidam com a correlação temporal através de algoritmos de estimação de movimento (GALL, 1991), que apresentam custo computacional elevado (CHAN; LEE, 1997).

Uma alternativa para evitar tais complexos algoritmos é a abordagem de codificação entre quadros (*interframe coding*), que aplica transformada de blocos em arranjos tridimensionais (NATARAJAN; AHMED, 1977; OHM; SCHAAR; WOODS, 2004; SAPONARA, 2012a).

Consequentemente, transformadas tridimensionais são levadas em conta. Codificação de vídeo baseada na DCT tridimensional (3D DCT) explora tanto a correlação espacial quanto a correlação temporal de *pixels*, uma vez que a propriedade de compactação da energia se estende adicionalmente para a dimensão temporal.

Em (BOZINOVIĆ; KONRAD, 2005; SERVAIS; JAGER, 1997; CHAN; LEE, 1997; MULLA et al., 2014; SAPONARA, 2012b; SAWANT; ADJEROH, 2011; BOZINOVIĆ; KONRAD, 2003; LAI; GUAN, 2002; CHAN; SIU, 1997; LEE; CHAN; ADJEROH, 1997), são propostos esquemas de compressão de vídeo baseados na divisão sucessiva de quadros em cubos de *pixels* que são então aplicados à 3D DCT. Em (CHAN; LEE, 1997), foi proposto um método para gerar um *volume* de quantização para coeficientes no domínio da frequência da 3D DCT que substitui as usuais matrizes de quantização, derivadas para coeficientes da 2D DCT. Recentemente, foi proposto o sistema para transmissão multi-usuário *SoftCast* (JAKUBCZAK; RAHUL; KATABI, 2010; JAKUBCZAK; KATABI, 2010; JAKUBCZAK; KATABI, 2011). Tal sistema emprega a 3D DCT, evitando algoritmos de compensação de movimento e codificação diferencial.

Em (LI; GUO; LU, 2016), as propriedades de decorrelação da 3D DCT são exploradas para um método de quantificação de qualidade de vídeo sem referência. A 3D DCT também foi empregada para obter características — dados que podem ser empregados para modelagem e classificação de sinais — para análise de movimento (BOZINOVIĆ; KONRAD, 2005), segmentação de imagem de fígado (DANCIU et al., 2012) e rastreamento visual (LI et al., 2013).

Adicionalmente, aplicações em tempo real de ultrassonografia médica e industrial requerem compressão de dados tridimensionais computacionalmente eficiente (GOVINDAN et al., 2016; DESMOULIERS; ORUKLU; SANIIE, 2009). Métodos baseados na 3D DCT são alternativas para este contexto (CHENG; SHEN; LI, 2012; ORUKLU; SANIIE, 2006; CARDOSO; SANIIE, 2004). Em (BOUSSAKTA, 2004), foi proposto o algoritmo 3D dizimação na frequência de base vetorial — ou *vector-radix decimation-in-frequency* (3D VR DIF) — que computa diretamente a 3D DCT, apresentando menor complexidade multiplicativa que o método empregando a decomposição linha-coluna.

Em (JACOB; KUMAR, 2015), foi apresentada uma implementação em arranjo de por-

tas programáveis em campo — ou *field-programmable gate array* (FPGA) — para a 3D DCT inteira para compressão de vídeo. Ainda mais, DCTs de dimensões mais altas, tais como a DCT quadridimensional (4D DCT), também apresentam aplicabilidade em diferentes contextos (DAI; CHEN; LIN, 2005), tais como renderização de campo de luz (KAMAL et al., 2016; CARVALHO et al., 2018; MAGNOR; GIROD, 2000; LEVOY; HANRAHAN, 1996), *lumi-graph* (GORTLER et al., 1996; ZHANG; LI, 2000) e codificação de vídeo (SANG et al., 2013; ANITHA; NAGABHUSHANA, 2011; XIANGWEN; HEXIN; ZHIJIE, 2004). Algoritmos rápidos para a DCT multidimensional foram propostos em (ZENG; BI; LEYMAN, 2000b; CHEN; DAI; LI, 2003; DAI; CHEN; LIN, 2005). Apesar de tal gama de aplicações, a proposição de aproximações para a 3D DCT como ferramentas de baixa complexidade ainda representam um campo pouco explorado, ao nosso melhor conhecimento. A 3D DCT inteira proposta em (JACOB; KUMAR, 2015) ainda requer um número grande de operações aritméticas, incluindo multiplicações. Ainda mais, até onde sabemos, aproximações para a 3D DCT carecem de uma formalização matemática.

Por sua vez, a DHT multidimensional (mD DHT) vem sendo empregada no contexto de compressão de imagens médicas (SUNDER; ESWARAN; SRIRAAM, 2006; DULEBA, 1999; SUNDER; ESWARAN; SRIRAAM, 2005; PAPITHA; NANCY; NEDUMARAN, 2013). Conforme mencionado anteriormente, alguns exemplos de transformadas discretas multidimensionais apresentam a propriedade de separabilidade de núcleo (GONZALEZ; WOODS, 2001), que permite o cômputo multidimensional empregando um algoritmo unidimensional de acordo com a decomposição linha-coluna (ZENG; BI; LEYMAN, 2000a; SONG; LI, 2013). Entretanto, a DHT multidimensional (mD DHT) não apresenta tal propriedade. Sendo assim, manipulações algébricas adicionais são necessárias para viabilizar o emprego da decomposição linha-coluna.

Em (SUNDER; ESWARAN; SRIRAAM, 2006), foi proposto um esquema que aplica a 3D DHT no contexto de codificação de imagens médicas. No entanto, algoritmos de baixa complexidade não são tratados em tal trabalho. Embora apresente vasto leque de aplicações, métodos aproximados de muito baixa complexidade para a DHT são pouco explorados na literatura. Apenas em (CINTRA, 2011; BOUGUEZEL; AHMAD; SWAMY, 2013) foram apresentados métodos para derivar aproximações para a DHT. Além disso, não há na literatura uma

abordagem de extensão de métodos aproximados para o caso da mD DHT, ao nosso melhor conhecimento. Particularmente, métodos de aproximação que levem em conta a propriedade de não-separabilidade da mD DHT exata devem ser estudados.

1.2 OBJETIVOS

Este trabalho tem como objetivo principal investigar e propor novos métodos de baixa complexidade computacional para aplicações de processamento de sinais multidimensionais. No Capítulo 2, visa-se derivar aproximações podadas de baixa complexidade para a DTT baseadas em uma aproximação do estado-da-arte que apresenta a mais baixa complexidade da literatura. Como resultado, espera-se uma redução ainda mais acentuada da complexidade computacional. Objetiva-se avaliar o desempenho dos métodos propostos no contexto de compressão de imagem e vídeo.

No Capítulo 3, visa-se propor um formalismo algébrico para derivar aproximações para a DCT multidimensional empregando teoria de tensores de alta ordem (LATHAUWER; MOOR, 1998; LATHAUWER; MOOR; VANDEWALLE, 2000). Busca-se derivar métodos aproximados de baixa complexidade para a 3D DCT de dimensão $8 \times 8 \times 8$ empregando o formalismo tensorial desenvolvido e diversas matrizes aproximadas para a DCT disponíveis na literatura. Além disso, objetiva-se realizar uma análise de complexidade computacional e uma análise de *trade-off*, comparando os métodos derivados com algoritmos usuais para a 3D DCT. Visa-se empregar os métodos tridimensionais de baixa complexidade propostos nos contextos de: (i) codificação de vídeo entre quadros (NATARAJAN; AHMED, 1977; MULLA et al., 2014; SAPONARA, 2012b; SAWANT; ADJEROH, 2011; BOZINOVIĆ; KONRAD, 2005; BOZINOVIĆ; KONRAD, 2003; LAI; GUAN, 2002; CHAN; LEE, 1997; CHAN; SIU, 1997; LEE; CHAN; ADJEROH, 1997; SERVAIS; JAGER, 1997) e (ii) rastreamento visual (LI et al., 2013; SMEULDERS et al., 2014; ROSS et al., 2008).

No Capítulo 4, busca-se derivar novas aproximações de baixa complexidade para a DHT de comprimento oito baseadas em representação por racionais diádicos. Além disso, tem-se como objetivo empregar as novas matrizes aproximadas propostas para derivar versões aproximadas para a 3D DHT empregando o conceito de tensores de alta ordem e levando em conta

as propriedades de não-separabilidade da DHT. Ainda mais, almeja-se empregar as aproximações tridimensionais para a DHT derivadas no contexto de compressão de imagens médicas, tais como como ressonância magnética (MR) e angiografia por raios-X (XA).

1.3 ESTRUTURA DO DOCUMENTO

O presente documento está organizado de acordo com o que se segue. A Seção 2.1 revisa conceitos matemáticos preliminares relacionados a momentos discretos, transformadas discretas multidimensionais, DTT e transformadas discretas aproximadas. Na Seção 2.2, aproximações podadas para a DTT são matematicamente desenvolvidas e a complexidade computacional é analisada. A Seção 2.3 emprega as DTTs aproximadas propostas em simulações de codificação de imagem e vídeo. Uma análise de desempenho é fornecida, comparando os métodos propostos com os métodos usuais. Na Seção 2.5, arquiteturas para as transformadas aproximadas são desenvolvidas em FPGA. Uma análise de recursos de *hardware* e energia é realizada, onde uma comparação com métodos usuais é fornecida. O trabalho de pesquisa do Capítulo 2 foi realizado em colaboração com os pesquisadores Dr. Fábio M. Bayer, Paulo A. M. Oliveira, Raíza S. Oliveira e Dr. Arjuna Madanayake.

Na Seção 3.1, a base matemática teórica é provida e o conceito de tensores de alta ordem é revisado. Na Seção 3.2, aproximações para a DCT multidimensional são endereçadas. O cômputo aproximado para os casos tridimensional e multidimensional são formalizados na Seção 3.2.1. A complexidade computacional é discutida na Seção 3.2.2. Uma análise de *trade-off* é fornecida na Seção 3.2.3. A Seção 3.3 aborda um experimento de codificação entre quadros de vídeo empregando as aproximações para a 3D DCT derivadas. Um método para modificar o volume de quantização de modo a explorar as propriedades de baixa complexidade das aproximações propostas também é apresentado. Como prova de conceito alternativa, na Seção 3.3.1, as aproximações propostas foram embutidas em um sistema de rastreamento visual originalmente baseado na 3D DCT.

A Seção 4.1 aborda a DHT e propõe aproximações para a DHT de comprimento oito. Na Seção 4.2, revisam-se os conceitos matemáticos necessários para a extensão das aproximações propostas para o caso tridimensional e diversos métodos aproximados são derivados para a

3D DHT de dimensões $8 \times 8 \times 8$. Na Seção 4.3, as aproximações para a 3D DHT derivadas são aplicadas em um esquema de compressão de imagens médicas. O Capítulo 3 e o Capítulo 4 foram desenvolvidos em colaboração com Dr. Fábio M. Bayer.

Conclusões e trabalhos futuros são descritos no Capítulo 5. No Apêndice A, são expostas as imagens empregadas nas simulações realizadas no Capítulo 2. No Apêndice B, são listadas as produções bibliográficas e trabalhos realizados e em andamento, derivados do presente trabalho de doutorado.

2 APROXIMAÇÕES PODADAS PARA A TRANSFORMADA DISCRETA DE TCHEBICHEF BIDIMENSIONAL

No contexto de processamento digital de sinais, momentos de imagem são funções que apresentam uma grande variedade de aplicações, tais como reconhecimento de padrão visual (LI et al., 2018; GORJI; HADDADNIA, 2015; FLUSSER; SUK, 1993; BELKASIM; SHRIDHAR; AHMADI, 1991), classificação de objeto (HEYWOOD; NOAKES, 1995; LIM et al., 2013), remoção de ruído (JI et al., 2009), detecção de borda (GESTER; SIMON, 2018; SUN et al., 2014; GHOSAL; MEHROTRA, 1993), sistemas de visão robótica (HUSSMANN; LIEPERT, 2009; MARKANDEY et al., 1992), marca d'água digital (ZHANG et al., 2011; BORS, 2006) e compressão de dados (MUKUNDAN, 2003; ABU et al., 2010; AL-HILO; GEORGE, 2008; HSU; TSAI, 1993). Particularmente, momentos baseados em funções ortogonais contínuas, tais como os polinômios de Legendre e Zernike, foram extensivamente estudados e aplicados em reconhecimento de padrão (CHONG; RAVEENDRAN; MUKUNDAN, 2004), marca d'água digital (ZHANG et al., 2011; KIM; LEE, 2003), indexação de imagem (MANDAL; ABOULNASR; PANCHANATHAN, 1996), reconhecimento facial (HADDADNIA; FAEZ; MOALLEM, 2001), detecção de características (KHOTANZAD; HONG, 1990). No entanto, momentos baseados em polinômios contínuos apresentam dificuldades relacionadas a aplicações em tempo real. Momentos bidimensionais baseados em polinômios ortogonais contínuos são definidos através de integral dupla, cujas implementações práticas envolvem aproximações por somatórios discretos, resultando em erros numéricos (YAP; PARAMESRAN, 2005). Além disso, a propriedade de ortogonalidade geralmente é válida apenas em um domínio específico, que geralmente difere do domínio de coordenadas da imagem (MUKUNDAN; ONG; LEE, 2001). Conseqüentemente, uma transformação de coordenadas é necessária, o que incrementa o custo computacional envolvido.

Uma alternativa para contornar tais problemas de implementação é empregar polinômios ortogonais discretos como funções de base (MUKUNDAN, 2003), tais como os polinômios de Tchebichef, Charlier, Meixner e Hahn (ABRAMOWITZ; STEGUN, 1970, p. 788). Uma vez que momentos baseados em polinômios discretos são inerentemente definidos através

de transformações discretas e a ortogonalidade é geralmente válida num domínio compatível com o domínio digital de coordenadas da imagem, suas implementações não requerem aproximação de integral nem transformação de coordenadas (MUKUNDAN; ONG; LEE, 2001).

Em particular, momentos discretos baseados no polinômio de Tchebichef têm sido considerados em aplicações tais como quantificação de artefato de bloco em imagem (LI et al., 2014), reconhecimento de imagem (ZHANG et al., 2010), verificação de integridade em imagens médicas (HUANG et al., 2012) e compressão de imagem (ISHWAR; MEHER; SWAMY, 2008; ABU et al., 2010; PRATTIPATI; SWAMY; MEHER, 2013; PRATTIPATI et al., 2013; SENAPATI; PATI; MAHAPATRA, 2014). Em aplicações de compressão de imagem, os momentos bidimensionais de Tchebichef são computados através de uma transformação bidimensional cujo núcleo são os polinômios de Tchebichef. Tal transformação é dada pela 2D DTT.

A DTT apresenta a propriedade de compactação de energia com desempenho em compressão de imagem e vídeo similar ao da DCT (HUNT; MUKUNDAN, 2004). De fato, funções cosseno podem ser expressas pelos polinômios de Tchebichef com argumentos trigonométricos (DUDGEON; MERSEREAU, 1990, p. 138). Particularmente, a DTT de comprimento oito é compatível com os diversos esquemas de compressão de imagem e vídeo mencionados anteriormente.

Entretanto, o cômputo da DTT requer várias operações aritméticas, o que pode limitar seu emprego em aplicações que exijam baixa potência e em tempo-real. Um algoritmo rápido para a 2D DTT de comprimento oito baseada numa representação em números inteiros foi proposta em (PRATTIPATI et al., 2013). Embora tal algoritmo reduza substancialmente o custo computacional, ele ainda apresenta um número consideravelmente alto de adições devido à representação CSD.

Levando em conta o leque de aplicações em que a DTT pode se inserir e ao fato de algoritmo rápidos neste caso serem pouco explorados, alternativas de baixa complexidade para tal transformada são desejáveis para aplicações práticas. Aproximações de baixa complexidade para a DTT com custo aditivo consideravelmente baixo foram propostas em (OLIVEIRA et al., 2015; OLIVEIRA et al., 2016). Tais aproximações permitem implementações de baixa complexidade por substituir os elementos dos vetores de base da DTT exata por elementos do

conjunto $\mathbb{P} = \{0, \pm 1, \pm \frac{1}{2}, \pm 2\}$ de acordo com um problema de otimização. Uma multiplicação por tais elementos pode ser implementada apenas com deslocamentos de *bits*, no máximo, sem custo aditivo extra. Paralelamente, foi proposto um esquema podado para a DTT exata que descarta o cômputo de elementos de alta frequência (KOUADRIA et al., 2017).

O objetivo deste capítulo é derivar versões podadas para a DTT aproximada em (OLIVEIRA et al., 2016) visando uma redução ainda mais acentuada de complexidade computacional em aplicações de codificação de imagem e vídeo.

2.1 REVISÃO MATEMÁTICA

Nesta seção, serão abordados conceitos básicos importantes para o desenvolvimento deste capítulo e do trabalho como um todo.

2.1.1 Momentos Baseados em Funções Contínuas e Discretas

De forma geral, o momento bidimensional de uma imagem 2D analógica $I_a(t_1, t_2)$ empregando uma função contínua $\psi_{pq}(t_1, t_2)$ é definido de acordo com o que se segue (MUKUNDAN; ONG; LEE, 2001):

$$M_{pq} \triangleq \int_x \int_y \psi_{pq}(t_1, t_2) I_a(t_1, t_2) dt_1 dt_2, \quad p, q = 0, 1, 2, \dots \quad (2.1)$$

Por exemplo, o momento bidimensional de Legendre é definido fazendo $\psi_{pq}(t_1, t_2) = P_p(t_1)P_q(t_2)$, $t_1, t_2 \in [-1, 1]$, em que

$$P_p(t) \triangleq \frac{1}{2^p p!} \frac{d^p}{dt^p} (t^2 - 1)^p$$

é o polinômio de Legendre de p -ésima ordem (LIAO; PAWLAK, 1996). De acordo com (MUKUNDAN; ONG; LEE, 2001; LIAO; PAWLAK, 1996), aproximações numéricas para (2.1) podem apresentar elevados erros e complexidades computacionais, especialmente para momentos de mais alta ordem.

Por outro lado, momentos baseados em polinômios discretos não dependem de aproximação numérica e possibilitam o desenvolvimento de algoritmos eficientes (MUKUNDAN; ONG; LEE, 2001; MUKUNDAN, 2003; HUNT; MUKUNDAN, 2004). O polinômio normali-

zado de Tchebichef de k -ésima ordem é dado por (OLIVEIRA et al., 2016):

$$\tilde{t}_k[n] \triangleq \sqrt{\frac{(2k+1)(N-k-1)!}{(N+k)!}} \cdot (1-N)_k \cdot {}_3F_2(-k, -n, 1+k; 1, 1-N; 1), \quad (2.2)$$

em que

$${}_pF_q(c_1, \dots, c_p; d_1, \dots, d_q; z) \triangleq \sum_{i=0}^{\infty} \frac{\prod_{r=1}^p (c_r)_i}{\prod_{r=1}^q (d_r)_i} \cdot \frac{z^i}{i!}$$

é a função hipergeométrica generalizada (GRINSHPAN, 2013), $(c)_i = c(c+1)(c+2)\cdots(c+i-1)$ é o fatorial ascendente e $n = 0, 1, \dots, N-1$. Os momentos bidimensionais de Tchebichef de uma imagem 2D discreta $x[n_1, n_2]$, para $n_1, n_2 = 0, 1, \dots, N-1$, são definidos empregando o polinômio mencionado anteriormente como núcleo, tal qual

$$T_{k_1 k_2} \triangleq \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} \tilde{t}_{k_1}(n_1) \tilde{t}_{k_2}(n_2) x[n_1, n_2], \quad k_1, k_2 = 0, 1, 2, \dots, N-1. \quad (2.3)$$

2.1.2 Transformadas Discreta Multidimensionais

O conceito de momento discreto está intrinsecamente relacionado ao de transformadas discretas. As definições a seguir são generalizações propostas tomando como referência os caso de transformadas discretas generalizadas de uma (OPPENHEIM; SCHAFER, 2010, p. 673) e duas (GONZALEZ; WOODS, 2001, p. 470) dimensões.

Definição 2.1. *Seja $x[n_1, n_2, \dots, n_R]$ um sinal R -dimensional, com $n_i = 0, 1, \dots, N_i - 1$, $i = 1, 2, \dots, R$. Uma transformada discreta R -dimensional é definida, de forma geral, pelo seguinte mapeamento:*

$$y[k_1, k_2, \dots, k_R] \triangleq \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \cdots \sum_{n_R=0}^{N_R-1} x[n_1, n_2, \dots, n_R] \Phi[n_1, k_1, n_2, k_2, \dots, n_R, k_R],$$

$$k_i = 0, 1, \dots, N_i - 1, \quad i = 1, 2, \dots, R, \quad (2.4)$$

em que $y[k_1, k_2, \dots, k_R]$ são os coeficientes da transformada, também chamado de sinal transformado, e $\Phi[n_1, k_1, n_2, k_2, \dots, n_R, k_R]$ é o **núcleo** da transformação ou **função de base** da transformada.

Para uma transformada R -dimensional de acordo com a Definição 2.1, a transformada inversa é dada na forma

$$x[n_1, n_2, \dots, n_R] = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \cdots \sum_{k_R=0}^{N_R-1} y[k_1, k_2, \dots, k_R] \Theta[n_1, k_1, n_2, k_2, \dots, n_R, k_R],$$

$$n_i = 0, 1, \dots, N_i - 1, \quad i = 1, 2, \dots, R,$$
(2.5)

em que $\Theta[n_1, k_1, n_2, k_2, \dots, n_R, k_R]$ é o núcleo ou função de base da transformada inversa.

Definição 2.2. Uma transformada discreta R -dimensional, dada conforme a Definição 2.1, apresenta a propriedade de **separabilidade** se seu núcleo de transformação pode ser expresso de acordo com:

$$\begin{aligned} \Phi[n_1, k_1, n_2, k_2, \dots, n_R, k_R] &= \phi_1[n_1, k_1] \cdot \phi_2[n_2, k_2] \cdots \phi_R[n_R, k_R] \\ &= \prod_{i=1}^R \phi_i[n_i, k_i]. \end{aligned}$$
(2.6)

Sendo assim, a transformada R -dimensional dada por (2.4) é dita **separável** e pode ser escrita de acordo com

$$y[k_1, k_2, \dots, k_R] = \sum_{n_1=0}^{N_1-1} \phi_1[n_1, k_1] \cdot \sum_{n_2=0}^{N_2-1} \phi_2[n_2, k_2] \cdots \sum_{n_R=0}^{N_R-1} \phi_R[n_R, k_R] \cdot x[n_1, n_2, \dots, n_R].$$
(2.7)

Uma transformada separável, de acordo com a Definição 2.2, é computada através de sucessivas transformações unidimensionais, como pode ser observado em (2.7).

Definição 2.3. Uma transformada R -dimensional separável, dada conforme a Definição 2.2, apresenta **núcleo regular**¹ se $N_1 = N_2 = \dots = N_R \triangleq N$ e os fatores do lado direito de (2.6) são funções iguais, isto é,

$$\phi_1[n, k] = \phi_2[n, k] = \dots = \phi_R[n, k] \triangleq \phi[n, k], \quad n, k = 0, 1, \dots, N - 1.$$

Sendo assim, o núcleo da transformada direta (2.4) é dado por

$$\Phi[n_1, k_1, n_2, k_2, \dots, n_R, k_R] = \prod_{i=1}^R \phi[n_i, k_i]$$
(2.8)

e a transformada é dita **separável e regular**.

¹ Em (GONZALEZ; WOODS, 2001), para o caso bidimensional, o autor define como *núcleo simétrico*. Entretanto, tal terminologia se confunde com a de *transformada simétrica*, que consiste em uma transformação linear cuja matriz de transformação é simétrica (BERNSTEIN, 2009, p. 165). Dessa forma, opta-se neste trabalho por terminologia que evite tal superposição de conceitos.

Uma transformada separável e regular, de acordo com a Definição 2.3, emprega a mesma transformação unidimensional sucessivamente. Dessa forma, um algoritmo rápido para a transformada unidimensional pode ser empregado para computar a respectiva versão R -dimensional separável e regular. Tal abordagem pode ser entendida como uma generalização da chamada *decomposição linha-coluna* (MADISETTI, 2009; ZENG; BI; LEYMAN, 2000a).

Definição 2.4. *Seja uma transformada R -dimensional separável com núcleo regular, de acordo com a Definição 2.3. A transformada é dita **ortogonal** se*

$$\sum_{n=0}^{N-1} \phi[n, k] \cdot \phi^*[n, \hat{k}] = C \delta_{k\hat{k}}, \quad (2.9)$$

em que C é uma constante², o superscrito $*$ indica complexo conjugado e

$$\delta_{k\hat{k}} \triangleq \begin{cases} 1, & \text{se } k = \hat{k}, \\ 0, & \text{caso contrário,} \end{cases} \quad (2.10)$$

é o delta de Kronecker (BERNSTEIN, 2009, p. xxi). Particularmente, a transformada é dita **ortonormal** ou **unitária** se $C = 1$.

Proposição 2.1. *Seja uma transformada R -dimensional separável com núcleo regular e ortogonal, de acordo com a Definição 2.4. Então, a transformada inversa, dada por (2.5), também é separável com núcleo regular e ortogonal. O núcleo da transformada inversa é dado por*

$$\Theta[n_1, k_1, \dots, n_R, k_R] = C^{-R} \prod_{i=1}^R \phi^*[n_i, k_i].$$

Demonstração. Primeiramente, note que qualquer sinal R -dimensional $y[k_1, \dots, k_R]$ pode ser expresso empregando o delta de Kronecker, dado por (2.10), de acordo com a seguinte identidade:

$$y[k_1, \dots, k_R] = \sum_{\hat{k}_1=0}^{N_1-1} \cdots \sum_{\hat{k}_R=0}^{N_R-1} y[\hat{k}_1, \dots, \hat{k}_R] \prod_{i=1}^R \delta_{k_i \hat{k}_i}. \quad (2.11)$$

² Em alguns casos, a constante C pode depender do índice k , portanto temos C_k . Entretanto, por simplicidade de notação e sem perda de generalidade, omite-se tal dependência.

A seguir, substituindo a expressão da transformada inversa (2.5) na expressão da transformada direta (2.4), tem-se

$$y[k_1, \dots, k_R] = \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_R=0}^{N_R-1} \left(\sum_{\hat{k}_1=0}^{N_1-1} \cdots \sum_{\hat{k}_R=0}^{N_R-1} y[\hat{k}_1, \dots, \hat{k}_R] \Theta[n_1, \hat{k}_1, \dots, n_R, \hat{k}_R] \right) \cdot \Phi[n_1, k_1, \dots, n_R, k_R]. \quad (2.12)$$

Reorganizando a ordem dos somatórios em (2.12), tem-se

$$y[k_1, \dots, k_R] = \sum_{\hat{k}_1=0}^{N_1-1} \cdots \sum_{\hat{k}_R=0}^{N_R-1} y[\hat{k}_1, \dots, \hat{k}_R] \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_R=0}^{N_R-1} \Phi[n_1, k_1, \dots, n_R, k_R] \cdot \Theta[n_1, \hat{k}_1, \dots, n_R, \hat{k}_R], \quad (2.13)$$

Comparando as expressões (2.13) e (2.11), devemos ter que

$$\sum_{n_1=0}^{N_1-1} \cdots \sum_{n_R=0}^{N_R-1} \Phi[n_1, k_1, \dots, n_R, k_R] \cdot \Theta[n_1, \hat{k}_1, \dots, n_R, \hat{k}_R] = \prod_{i=1}^R \delta_{k_i \hat{k}_i}. \quad (2.14)$$

Por suposição, a transformada é separável e de núcleo simétrico, portanto (2.8) é válida. Substituindo (2.8) em (2.14), tem-se

$$\sum_{n_1=0}^{N_1-1} \cdots \sum_{n_R=0}^{N_R-1} \left(\prod_{i=1}^R \phi[n_i, k_i] \right) \cdot \Theta[n_1, \hat{k}_1, \dots, n_R, \hat{k}_R] = \prod_{i=1}^R \delta_{k_i \hat{k}_i}. \quad (2.15)$$

Também por suposição, a transformada é ortogonal, portanto a expressão (2.9) também é válida.

Multiplicando ambos os lados de (2.9) para as R dimensões da transformada, verifica-se que

$$\left(\sum_{n_1=0}^{N_1-1} \phi[n_1, k_1] \cdot \phi^*[n_1, \hat{k}_1] \right) \cdots \left(\sum_{n_R=0}^{N_R-1} \phi[n_R, k_R] \cdot \phi^*[n_R, \hat{k}_R] \right) = (C \delta_{k_1 \hat{k}_1}) \cdots (C \delta_{k_R \hat{k}_R}). \quad (2.16)$$

Reorganizando (2.16), temos após algumas manipulações

$$\sum_{n_1=0}^{N_1-1} \cdots \sum_{n_R=0}^{N_R-1} \left(\prod_{i=1}^R \phi[n_i, k_i] \right) \cdot \left(C^{-R} \prod_{i=1}^R \phi^*[n_i, \hat{k}_i] \right) = \prod_{i=1}^R \delta_{k_i \hat{k}_i}. \quad (2.17)$$

Comparando (2.15) e (2.17), conclui-se que

$$\Theta[n_1, k_1, \dots, n_R, k_R] = C^{-R} \prod_{i=1}^R \phi^*[n_i, k_i].$$

□

2.1.3 Transformada Discreta de Tchebichef

De acordo com a Definição 2.1, verifica-se que a expressão para os momentos bidimensionais de Tchebichef (2.3) consiste em uma transformada bidimensional discreta ($R = 2$). Tal transformada é propriamente a 2D DTT (NAKAGAKI; MUKUNDAN, 2007). Além disso, o núcleo da 2D DTT é separável e regular, de acordo com a Definição 2.3, dado por $\Phi[n_1, k_1, n_2, k_2] = \tilde{t}_{k_1}[n_1] \cdot \tilde{t}_{k_2}[n_2]$. Adicionalmente, os polinômios normalizados de Tchebichef (2.2) representam um conjunto de funções reais e ortonormais (ABRAMOWITZ; STEGUN, 1970; ISHWAR; MEHER; SWAMY, 2008; MUKUNDAN; ONG; LEE, 2001; MUKUNDAN, 2003; OLIVEIRA et al., 2016), isto é

$$\sum_{n=0}^{N-1} \tilde{t}_k[n] \cdot \tilde{t}_{\hat{k}}[n] = \delta_{k\hat{k}}.$$

Consequentemente, a 2D DTT em (2.3) é ortonormal de acordo com a Definição 2.4.

Frequentemente, é conveniente usar a notação matricial para transformadas discretas. Considere uma DTT unidimensional definida com $R = 1$ em (2.4) e com núcleo $\Phi[n, k] = \tilde{t}_k[n]$. Sendo assim, a DTT de um sinal discreto unidimensional, representado pelo vetor coluna $\mathbf{x} = \begin{bmatrix} x[0] & x[1] & \dots & x[N-1] \end{bmatrix}^\top$, pode ser expressa através do vetor coluna de saída $\mathbf{y} = \begin{bmatrix} y[0] & y[1] & \dots & y[N-1] \end{bmatrix}^\top$, dado por

$$\mathbf{y} = \mathbf{T} \cdot \mathbf{x},$$

em que \mathbf{T} é a matriz de transformação da DTT, cujas entradas são dadas por

$$\left[\mathbf{T} \right]_{k,n} = \tilde{t}_k[n], \quad k, n = 0, 1, \dots, N-1.$$

Devido a ortonormalidade dos polinômios de Tchebichef, tem-se que $\mathbf{T} \cdot \mathbf{T}^\top = \mathbf{I}_N$, em que \mathbf{I}_N é a matriz identidade de ordem N . Logo, tem-se que $\mathbf{T}^{-1} = \mathbf{T}^\top$. Portanto, a transformada inversa é dada por $\mathbf{x} = \mathbf{T}^\top \cdot \mathbf{y}$.

De forma análoga, a 2D DTT em (2.3) pode ser representada por uma matriz de entrada $[\mathbf{X}]_{n_1, n_2} = x[n_1, n_2]$ com $n_1, n_2 = 0, 1, \dots, N-1$ e a matriz de saída $[\mathbf{Y}]_{k_1, k_2} = y[k_1, k_2]$ com $k_1, k_2 = 0, 1, \dots, N-1$, relacionadas pelas seguintes expressões direta e inversa, respec-

tivamente:

$$\mathbf{Y} = \mathbf{T} \cdot \mathbf{X} \cdot \mathbf{T}^\top,$$

$$\mathbf{X} = \mathbf{T}^\top \cdot \mathbf{Y} \cdot \mathbf{T}.$$

2.1.4 Aproximações para Transformadas Discretas e Medidas de Desempenho

Uma aproximação para uma transformada discreta objetiva ter um comportamento similar à sua versão exata de acordo com algum critério (CINTRA; BAYER; TABLADA, 2014). Seja uma transformada discreta de comprimento N , representada pela matriz de transformação \mathbf{T} de dimensão $N \times N$. Uma transformada aproximada é representada por uma matriz $\hat{\mathbf{T}}$ de mesma dimensão e que preserva propriedades úteis de uma dada matriz de transformação \mathbf{T} .

Uma transformada aproximada unidimensional de um dado vetor coluna \mathbf{x} é o vetor transformado \mathbf{y} , definido conforme se segue:

$$\mathbf{y} \triangleq \hat{\mathbf{T}} \cdot \mathbf{x}. \quad (2.18)$$

Definição 2.5. *Seja uma transformada representada pela matriz $\hat{\mathbf{T}}$, de dimensão $N \times N$, com entradas definidas sobre o conjunto \mathbb{F} , em que \mathbb{F} pode ser tanto o conjunto dos números complexos \mathbb{C} quanto o conjunto dos números reais \mathbb{R} . Seja*

$$\mathbf{D} = \left(\hat{\mathbf{T}} \cdot \hat{\mathbf{T}}^* \right)^{-1}, \quad (2.19)$$

em que o superscrito $*$ representa conjugado transposto (BERNSTEIN, 2009, p. xxiv). Se \mathbf{D} é uma matriz diagonal, a transformada $\hat{\mathbf{T}}$ é dita **ortogonal**. Particularmente, se \mathbf{D} resulta na matriz identidade \mathbf{I}_N , então $\hat{\mathbf{T}}$ é dita **ortonormal** ou **unitária** (CINTRA; BAYER; TABLADA, 2014).

Sendo a matriz $\hat{\mathbf{T}}$ ortogonal e definida sobre $\mathbb{F} = \mathbb{R}$, tem-se que $\hat{\mathbf{T}}^{-1} = \hat{\mathbf{T}}^* \cdot \mathbf{D} = \hat{\mathbf{T}}^\top \cdot \mathbf{D}$ e a transformada inversa de (2.18) é dada por

$$\mathbf{x} = \hat{\mathbf{T}}^\top \cdot \mathbf{D} \cdot \mathbf{y}.$$

No contexto de compressão de imagem, matrizes diagonais não introduzem sobrecarga computacional uma vez que podem ser incorporadas às etapas de quantização e desquantização de padrões de codificação de imagem e vídeo (BRITANAK; YIP; RAO, 2007).

Uma matriz ortogonal pode ser ortonormalizada pelo processo de decomposição polar (BERNSTEIN, 2009, p. 359), em que a transformada ortonormal é dada por

$$\tilde{\mathbf{T}} = \hat{\mathbf{S}} \cdot \hat{\mathbf{T}},$$

em que $\hat{\mathbf{T}}$ é uma transformada ortogonal e $\hat{\mathbf{S}} = \sqrt{\mathbf{D}}$, $\sqrt{\cdot}$ retorna a raiz quadrada matricial (CINTRA; BAYER; TABLADA, 2014). Consequentemente, $\tilde{\mathbf{T}} \cdot \tilde{\mathbf{T}}^\top = \mathbf{I}$.

Transformadas exatas e aproximadas são usualmente comparadas por métricas quantitativas tais como o erro médio quadrático (MSE) (BRITANAK; YIP; RAO, 2007, p. 162) e o ganho de codificação C_g (BRITANAK; YIP; RAO, 2007, p. 163). O MSE representa a energia do sinal de erro entre os sinais transformados exato e aproximado. Seja \mathbf{T} a matriz da transformada original, $\hat{\mathbf{T}}$ a matriz da transformada aproximada e $\mathbf{\Gamma} \triangleq \mathbf{T} - \hat{\mathbf{T}}$. O MSE é definido como

$$\text{MSE} \triangleq \frac{1}{N} \text{tr}(\mathbf{\Gamma} \cdot \mathbf{R}_x \cdot \mathbf{\Gamma}^\top), \quad (2.20)$$

em que \mathbf{R}_x é a matriz de covariância do vetor de entrada \mathbf{x} e $\text{tr}(\cdot)$ representa o traço de uma matriz. No contexto de compressão de imagem, geralmente se considera a matriz de correlação para um processo markoviano de primeira ordem com coeficiente de correlação $\rho = 0.95$. Uma boa aproximação tende a apresentar baixos valores para MSE—em geral, da ordem de 10^{-2} .

O ganho de codificação (*transform coding gain*) C_g de uma transformada é uma métrica para o desempenho em codificação em aplicações de compressão de dados. Seja uma matriz de transformação $\hat{\mathbf{T}}$ de dimensão $N \times N$; \mathbf{h}_k e \mathbf{g}_k os vetores linha de $\hat{\mathbf{C}}$ e $\hat{\mathbf{C}}^{-1}$, respectivamente, para $k = 0, 1, \dots, N - 1$; o ganho de codificação é definido como (BRITANAK; YIP; RAO, 2007)

$$C_g \triangleq 10 \log_{10} \left(\prod_{k=0}^{N-1} \frac{1}{(A_k \cdot B_k)^{\frac{1}{N}}} \right), \quad (2.21)$$

em que $A_k = \text{sum}[(\mathbf{h}_k^\top \cdot \mathbf{h}_k) \odot \mathbf{R}_x]$, retorna a soma dos elementos da matriz de seu argumento, o operador \odot representa o produto matricial de Hadamard (CAO et al., 2013) — produto elemento a elemento — e $B_k = \|\mathbf{g}_k\|^2$, em que $\|\cdot\|$ representa a norma euclidiana.

Além disso, aproximações são tipicamente construídas para apresentar propriedades de baixa complexidade. Em geral, a transformada exata apresenta matriz de transformação com coeficientes no campo dos reais ou complexos, enquanto as aproximações são derivadas com valores numéricos selecionados, tais como os racionais diádicos (BRITANAK; YIP; RAO, 2007, p. 181), que são adequados para implementações de baixa complexidade (GEREK; ÇETIN, 2006).

Definição 2.6. *Um racional diádico é um número $d \in \mathbb{Q}$ que pode ser representado na forma*

$$d = \frac{q}{2^p},$$

em que $q, p \in \mathbb{Z}$, \mathbb{Q} é o conjunto dos números racionais e \mathbb{Z} o conjunto dos números inteiros.

Racionais diádicos podem ser expandidos considerando a representação canônica sinalizada (*canonic signed representation*) (CSD) (LEE, 1997), conforme a expressão:

$$d = \sum_i s_i \cdot 2^{-p_i}, \quad (2.22)$$

em que $s_i \in \{0, \pm 1\}$ e $p_i \in \{0, 1, \dots, p\}$. Dessa forma, o produto de um inteiro arbitrário por um racional diádico pode ser implementado com um conjunto de operações de adição e deslocamento de *bits*, proporcionando economia em termos de *hardware*, tempo e energia quando comparado com um algoritmo de multiplicação de inteiros usual (BOUGUEZEL; AHMAD; SWAMY, 2008; BOSSEN et al., 2012; OLIVEIRA et al., 2017). O conjunto $\mathbb{P} = \{0, \pm 1, \pm 2, \pm \frac{1}{2}\}$ é um subconjunto dos conjunto dos racionais diádicos.

Aproximações não ortogonais também são encontradas na literatura (HAWHEEL, 2001; CINTRA; BAYER; TABLADA, 2014; CINTRA, 2011). Em tal contexto, a transformada aproximada pode apresentar propriedades de *quase-ortogonalidade*, em que o produto $\hat{\mathbf{T}}^\top \cdot \hat{\mathbf{T}}$ é “quase” uma matriz diagonal. A métrica *desvio de diagonalidade* $\partial \left(\hat{\mathbf{T}}^\top \cdot \hat{\mathbf{T}} \right)$ foi proposta por Cintra et al. em (CINTRA; BAYER; TABLADA, 2014), a qual permite quantificar a quase-ortogonalidade de uma dada aproximação não-ortogonal. O desvio de diagonalidade de uma matriz quadrada \mathbf{M} qualquer é definido de acordo com (CINTRA; BAYER; TABLADA, 2014)

$$\partial(\mathbf{M}) \triangleq 1 - \frac{\|\text{diag}(\mathbf{M})\|_F}{\|\mathbf{M}\|_F}, \quad (2.23)$$

em que \mathbf{M} é uma matriz quadrada genérica, $\|\cdot\|_F$ é a norma matricial de Frobenius (BERNSTEIN, 2009, p. 547) e $\text{diag}(\cdot)$ retorna uma matriz diagonal cujos elementos são os elementos diagonais da matriz de seu argumento.

Se $\partial(\mathbf{M}) = 0$, então \mathbf{M} é uma matriz diagonal. Quão menor for o valor de $\partial(\mathbf{M})$, mais próxima será \mathbf{M} de uma matriz diagonal. Valores mais altos apresentam uma interpretação oposta. No contexto de compressão de imagem, o desvio de diagonalidade de 0.1056 associado à SDCT (HAWHEEL, 2001) — aproximação clássica para a DCT de comprimento oito — é frequentemente adotado como limiar de referência para quase-ortogonalidade (OLIVEIRA et al., 2016; CINTRA; BAYER; TABLADA, 2014). O desenvolvimento matemático para o caso ortogonal descrito anteriormente pode ser generalizado para aproximações quase-ortogonais: $\mathbf{D}_N = \text{diag}\left(\left(\hat{\mathbf{T}}_N \cdot \hat{\mathbf{T}}_N^\top\right)^{-1}\right)$ (CINTRA; BAYER; TABLADA, 2014).

2.1.5 Uma Aproximação de Baixa Complexidade para a DTT

Uma aproximação de baixa complexidade para a DTT de comprimento oito foi proposta em (OLIVEIRA et al., 2016). Tal aproximação $\tilde{\mathbf{T}}$ é dada por:

$$\tilde{\mathbf{T}} = \hat{\mathbf{S}} \cdot \hat{\mathbf{T}}, \quad (2.24)$$

em que

$$\hat{\mathbf{T}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & -1 & 0 & 0 & 1 & 1 & 2 \\ 2 & 0 & -1 & -1 & -1 & -1 & 0 & 2 \\ -2 & 1 & 2 & 1 & -1 & -2 & -1 & 2 \\ 1 & -2 & 0 & 1 & 1 & 0 & -2 & 1 \\ -1 & 2 & -1 & -1 & 1 & 1 & -2 & 1 \\ 0 & -1 & 2 & -1 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -2 & 1 & 0 & 0 \end{bmatrix}$$

é uma matriz de baixa complexidade, $\hat{\mathbf{S}} = \sqrt{\left[\text{diag}\left(\hat{\mathbf{T}} \cdot \hat{\mathbf{T}}^\top\right)\right]^{-1}}$ é a matriz diagonal para normalização dos vetores de base (CINTRA; BAYER; TABLADA, 2014).

A matriz aproximada dada por $\tilde{\mathbf{T}}$ em (2.24) não é ortogonal e apresenta desvio de diagonalidade $\partial \left(\hat{\mathbf{T}}^\top \cdot \hat{\mathbf{T}} \right) = 0.0453$. Portanto, a matriz $\hat{\mathbf{T}}$ pode ser considerada quase-ortogonal. Dessa forma, considera-se $\tilde{\mathbf{T}}^{-1} \approx \tilde{\mathbf{T}}^\top$. As transformações bidimensionais direta e inversa são respectivamente dadas por:

$$\begin{aligned} \mathbf{Y} &= \tilde{\mathbf{T}} \cdot \mathbf{X} \cdot \tilde{\mathbf{T}}^\top, \\ \mathbf{X} &\approx \tilde{\mathbf{T}}^\top \cdot \mathbf{Y} \cdot \tilde{\mathbf{T}}. \end{aligned} \tag{2.25}$$

2.2 APROXIMAÇÕES PODADAS PROPOSTAS

Na presente seção, são propostas aproximações podadas para a DTT. As novas transformadas são derivadas algebricamente e suas complexidades são analisadas.

2.2.1 Definição Matemática

A abordagem de transformadas podadas (*transform pruning*) pode ser modelado matematicamente através da exclusão de linhas de mais alto índice na matriz de transformação original (CINTRA et al., 2016; KOUADRIA et al., 2017). Seja $\hat{\mathbf{T}}$ uma matriz de dimensão $N \times N$. Então, uma nova matriz de transformação com *parâmetro de poda* K é formalizada de acordo com o seguinte:

$$\tilde{\mathbf{T}}_{\langle K \rangle} = \hat{\mathbf{S}}_K \cdot \hat{\mathbf{T}}_{\langle K \rangle}, \tag{2.26}$$

em que $\hat{\mathbf{T}}_{\langle K \rangle}$ é definida como uma matriz de dimensão $K \times N$ com as primeiras K linhas de $\hat{\mathbf{T}}$ e $\hat{\mathbf{S}}_K = \sqrt{\left[\text{diag} \left(\hat{\mathbf{T}}_{\langle K \rangle} \cdot \hat{\mathbf{T}}_{\langle K \rangle}^\top \right) \right]^{-1}}$ (CINTRA et al., 2016; CINTRA; BAYER; TABLADA, 2014). A 2D DTT aproximada direta é dada pela expressão

$$\tilde{\mathbf{Y}}_K = \tilde{\mathbf{T}}_{\langle K \rangle} \cdot \mathbf{X} \cdot \tilde{\mathbf{T}}_{\langle K \rangle}^\top. \tag{2.27}$$

Note que a matriz de saída $\tilde{\mathbf{Y}}_K$, de dimensão $K \times K$, apresenta apenas os K^2 coeficientes da parte superior a esquerda da matriz $\tilde{\mathbf{Y}}$, que geralmente retêm a maior parte da energia do sinal transformado (LECUIRE; MAKKAOUI; MOUREAUX, 2012; CINTRA et al., 2016).

A nova matriz de transformação $\tilde{\mathbf{T}}_{\langle K \rangle}$ não é quadrada, apresentando dimensão $K \times N$. Consequentemente, a invertibilidade usual não é admitida (BERNSTEIN, 2009, p. 98). Neste

contexto, podem ser considerados métodos de inversas generalizadas, tal qual a inversa generalizada de Moore-Penrose (BERNSTEIN, 2009, p. 363). A inversa generalizada de uma matriz $\tilde{\mathbf{T}}_{\langle K \rangle}$ é denotada por $\tilde{\mathbf{T}}_{\langle K \rangle}^+$. O produto $\tilde{\mathbf{T}}_{\langle K \rangle}^+ \cdot \tilde{\mathbf{T}}_{\langle K \rangle}$ geralmente não resulta em uma matriz diagonal. Sendo assim, um desvio de diagonalidade é esperado uma vez que se considera tal método de inversa generalizada.

Diversos métodos podados aproximados propostos para a DCT que foram derivados de aproximações ortogonais apresentam que a inversa generalizada da matriz podada é idêntica a matriz podada transposta propriamente (CINTRA et al., 2016; COUTINHO et al., 2015). Neste caso, a inversa generalizada herda as características de baixa complexidade da matriz aproximada podada. Entretanto, as matrizes podadas aproximadas propostas para a DTT não demonstram esta propriedade e as inversas generalizadas obtidas não apresentam atributos de baixo custo computacional.

Para explorar as propriedades de baixo custo computacional das matrizes podadas aproximadas propostas para a transformação direta, adota-se a matriz transposta $\tilde{\mathbf{T}}_{\langle K \rangle}^\top$ como uma aproximação para a inversa generalizada $\tilde{\mathbf{T}}_{\langle K \rangle}^+$ no cálculo da transformada inversa podada. Para avaliar quantitativamente tal abordagem proposta, verificou-se o desvio de diagonalidade para ambos os casos. Tal comparação é mostrada na Tabela 1. Note que os desvios de diagonalidade para ambas as abordagens são similares, sendo o desvio de diagonalidade associado à matriz transposta ainda menor do que o da inversa generalizada. Na Tabela 1, mostra-se também a redução percentual do desvio de diagonalidade do método transposto comparado com o método de inversa generalizada. Tal resultado sugere que as matrizes podadas transpostas para a DTT aproximada são mais apropriadas para o cômputo da transformada inversa do que as matrizes obtidas pela inversa generalizada de Moore-Penrose de acordo com ambos os critérios de desvio de diagonalidade e de custo computacional. Consequentemente, a transformada inversa podada é computada de acordo com

$$\mathbf{X} \approx \tilde{\mathbf{T}}_{\langle K \rangle}^\top \cdot \mathbf{Y}_K \cdot \tilde{\mathbf{T}}_{\langle K \rangle}. \quad (2.28)$$

Tabela 1 – Desvios de diagonalidade associados às inversas generalizadas e às matrizes transpostas e a diferença percentual para cada 2D DTT aproximada podada com parâmetro de poda K

K	$\partial \left(\tilde{\mathbf{T}}_{(K)}^+ \cdot \tilde{\mathbf{T}}_{(K)} \right)$	$\partial \left(\tilde{\mathbf{T}}_{(K)}^\top \cdot \tilde{\mathbf{T}}_{(K)} \right)$	Reduction (%)
7	0.0525	0.0494	5.6 %
6	0.0972	0.0853	12.2 %
5	0.1627	0.1420	12.7 %
4	0.1951	0.1751	10.3 %
3	0.2703	0.2654	1.8 %
2	0.4410	0.4367	1.0 %

Fonte: (COUTINHO et al., 2018)

2.2.2 Análise de Complexidade

As aproximações podadas para a 2D DTT são definidas preservando a propriedade de separabilidade, de acordo com (2.25). Sendo assim as transformações 2D podem ser computadas de acordo com o método de decomposição linha-coluna (ZENG; BI; LEYMAN, 2000a). Nesse caso e para uma transformada podada, um algoritmo 1D é aplicado N vezes para cada linha da matriz de entrada, resultando em uma matriz intermediária de dimensão $N \times K$. Então, o algoritmo 1D é aplicado novamente a cada uma das K colunas resultantes da matriz intermediária. Através de tal decomposição linha-coluna, a complexidade aritmética bidimensional C_{2D} se relaciona com a complexidade aritmética unidimensional C_{1D} de acordo com a seguinte expressão (CINTRA et al., 2016):

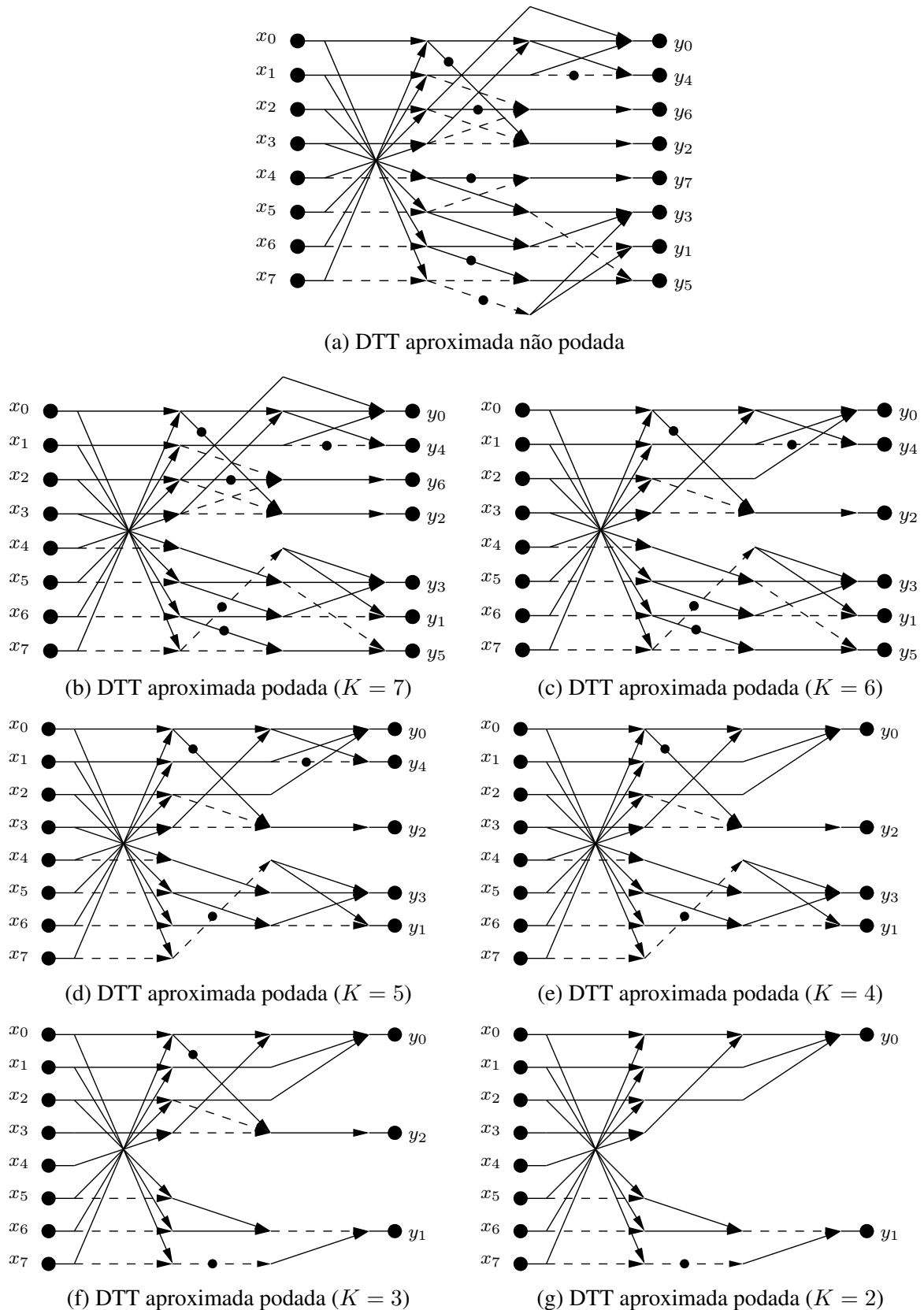
$$C_{2D} = (K + N) \cdot C_{1D}, \quad (2.29)$$

em que K é o parâmetro para a aproximação podada em (2.26). Note que, para o caso usual não-podado, $K = N$ e então $C_{2D} = 2N \cdot C_{1D}$. A Equação (2.29) deixa claro que um método podado, i.e $K < N$, é vantajoso para o caso bidimensional por reduzir ambos os termos $(K + N)$ e C_{1D} .

A Figura 1 mostra os algoritmos unidimensionais derivados a partir das matrizes podadas aproximadas propostas para a DTT, os quais são obtidos do algoritmo original eliminando as linhas de fluxo de sinal que se tornam desnecessárias ao eliminar coeficientes de saída (OLIVEIRA et al., 2016). A Figura 2 mostra, em vermelho, as linhas de fluxo do algoritmo original que foram podadas para obter o novo o algoritmo para cada parâmetro de poda K . A Tabela 2

mostra a complexidade aritmética bidimensional para todos os métodos propostos, obtidas aplicando (2.29) e fazendo $K \in [2, 7]$. Para comparação, também está incluída a complexidade da 2D DTT de dimensão 8×8 a partir do algoritmo proposto em (PRATTIPATI et al., 2013). Para $K = 1$, o método podado retorna apenas o termo DC da transformada, caso que já foi estudado previamente para algoritmos podados para a DCT (CINTRA et al., 2016).

Figura 1 – Algoritmos rápidos para os aproximações podadas para a DTT propostas para cada parâmetro de poda K . O círculo preto \bullet representa um deslocamento de *bits* para a esquerda (multiplicação por 2).



Fonte: (COUTINHO et al., 2018)

Figura 2 – Algoritmos rápidos para os aproximações podadas para a DTT propostas para cada parâmetro de poda K mostrando as linhas de fluxos podadas, em vermelho.

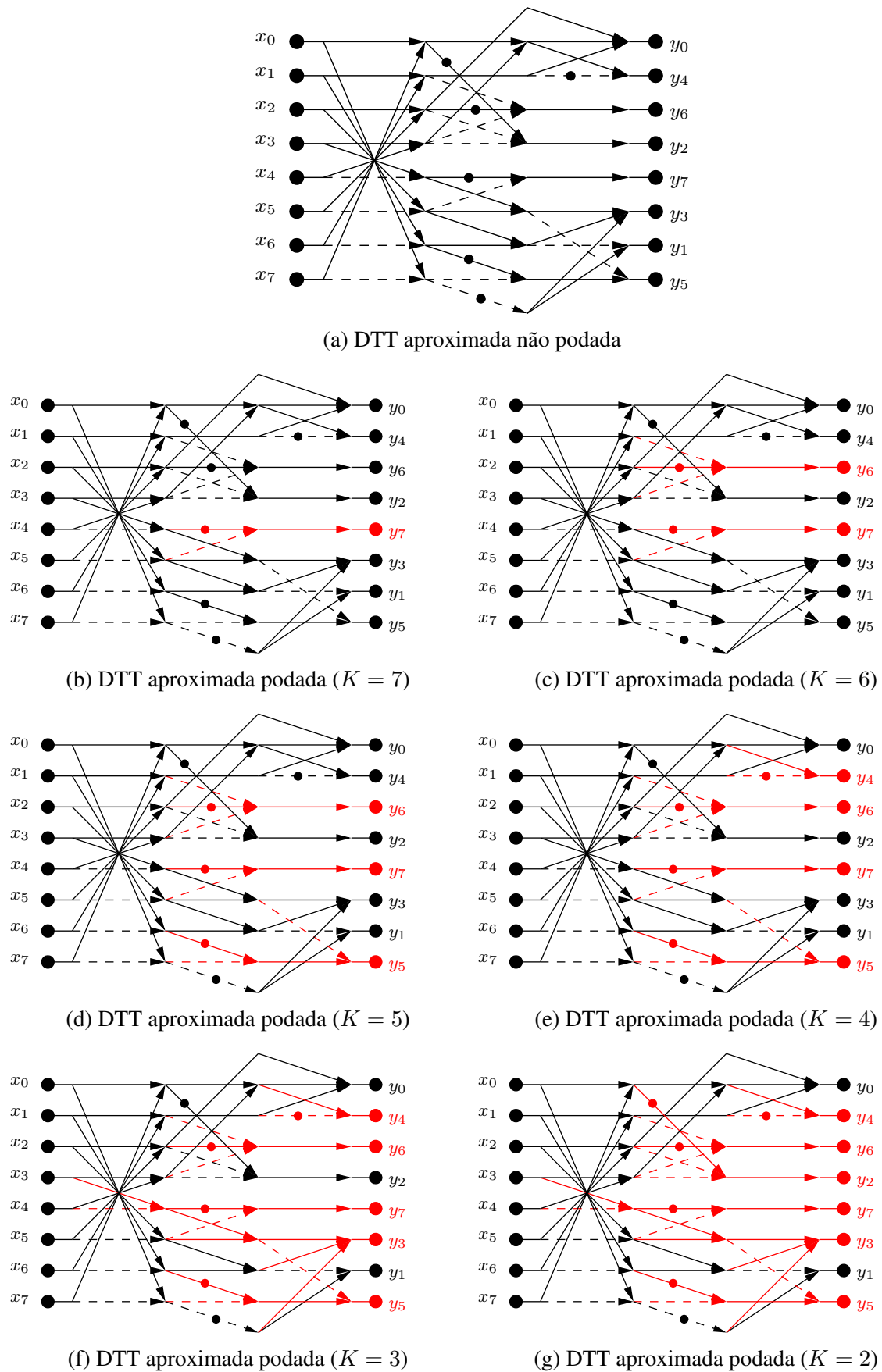


Tabela 2 – Complexidade das aproximações original e podadas para a 2D DTT de dimensões 8×8 e redução percentual relativa à 2D DTT exata

Método	Complexidade 2D		Redução (%)	
	Adições	Desloc. de <i>bits</i>	Adições	Desloc. de <i>bits</i>
2D DTT exata \mathbf{T}	704	464	—	—
Aprox. original $\hat{\mathbf{T}}$	384	96	45.45	79.31
Aprox. podada $\hat{\mathbf{T}}_{(7)}$	345	75	50.99	83.84
Aprox. podada $\hat{\mathbf{T}}_{(6)}$	294	56	58.24	87.93
Aprox. podada $\hat{\mathbf{T}}_{(5)}$	247	39	64.91	91.59
Aprox. podada $\hat{\mathbf{T}}_{(4)}$	216	24	69.32	94.83
Aprox. podada $\hat{\mathbf{T}}_{(3)}$	154	22	78.12	95.26
Aprox. podada $\hat{\mathbf{T}}_{(2)}$	120	10	82.95	97.84

Fonte: (COUTINHO et al., 2018)

A Tabela 2 mostra que as aproximações podadas propostas para a 2D DTT apresentam uma redução relevante no número de operações aritméticas. Nas seções seguintes, estudaremos o desempenho de tais métodos propostos no contexto de compressão da imagem e vídeo.

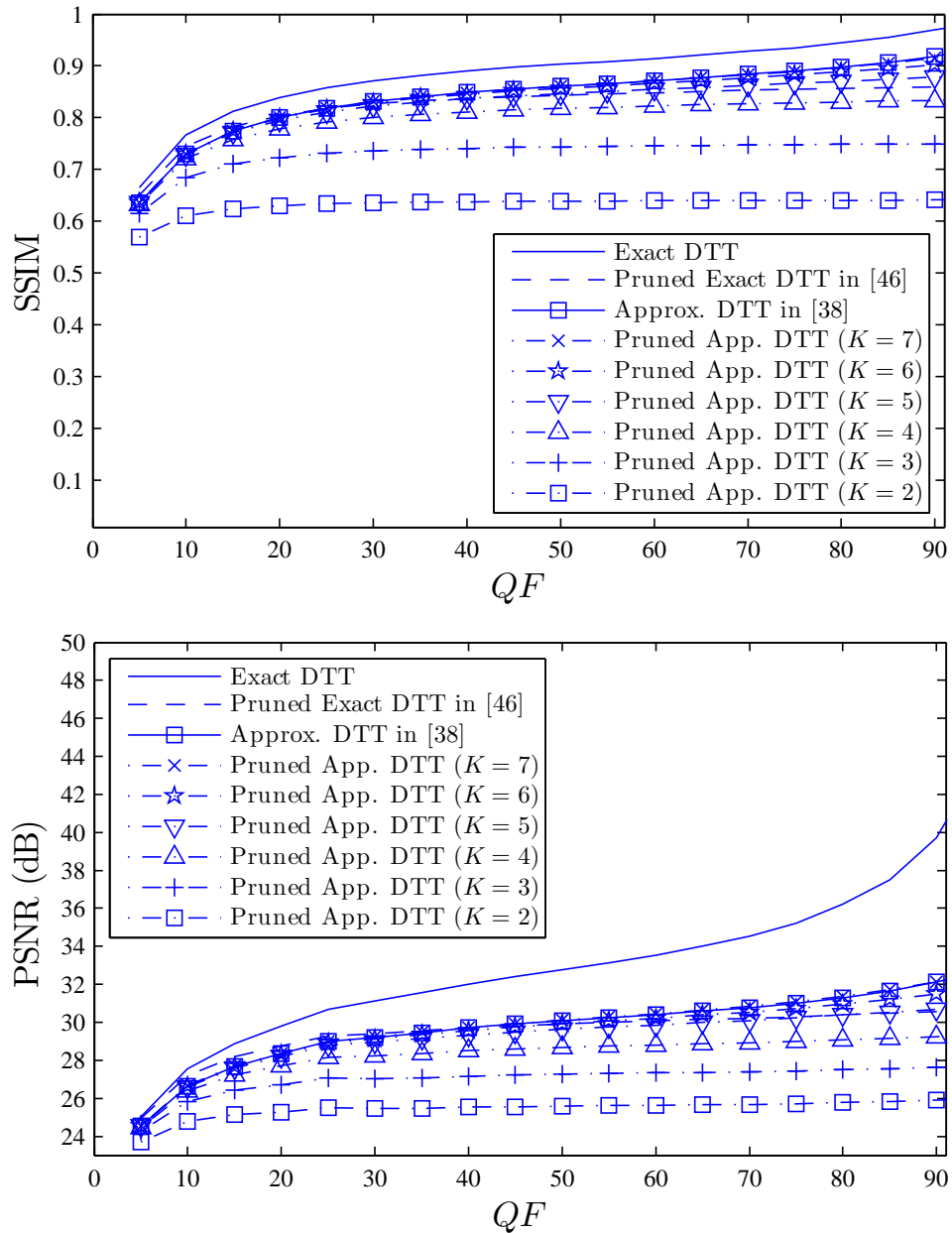
2.3 APLICAÇÃO EM CODIFICAÇÃO DE IMAGEM E VÍDEO

Na presente seção, as aproximações podadas propostas são consideradas em esquemas de codificação de imagem e vídeo. Uma simulação baseada no algoritmo de compressão de imagem JPEG é realizada com os métodos propostos. Além disso, os métodos propostos foram empregados nos padrões de codificação de vídeo: H.264/AVC (TEAM, 2017) e H.265/HEVC (International Telecommunication Union, 2013).

2.3.1 Compressão de Imagem Tipo JPEG

Todos os métodos podados derivados da matriz aproximada (2.26) foram submetidos a um esquema de compressão baseado no padrão JPEG (WALLACE, 1992). Foi considerado um conjunto de cinquenta imagens disponíveis na base de dados pública disponível em (Signal and Image Processing Institute, 2018). Cada imagem foi dividida em um conjunto de blocos de *pixels* de dimensões 8×8 e cada bloco \mathbf{X} foi então submetido à transformação dada por (2.27). Para comparação, também foram consideradas as 2D DTT exata e 2D DTT aproximada original não podada. Então, cada bloco transformado obtido \mathbf{Y} foi quantizado.

Figura 3 – Métricas de desempenho em simulação tipo JPEG para a 2D DTT exata, a 2D DTT aproximada e as 2D DTT aproximadas podadas para cada fator de poda K .



Fonte: (COUTINHO et al., 2018)

O processo de quantização foi realizado de acordo com a divisão elemento por elemento das entradas do bloco transformado \mathbf{Y} pelas entradas da matriz de quantização \mathbf{Q} . A matriz de quantização \mathbf{Q} é dada por (OLIVEIRA et al., 2016) $\mathbf{Q} = \lfloor (S \cdot \mathbf{Q}_0 + 50) / 100 \rfloor$, em que $S = 5000/QF$, se $QF < 50$, e $200 - 2 \cdot QF$ caso contrário, $QF \in (0, 100]$ é um fator de qualidade e \mathbf{Q}_0 é a matriz de quantização padrão para luminância dada por (BHASKARAN; KONSTANTINIDES, 1997, p. 81):

$$\mathbf{Q}_0 = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}.$$

Para valores baixos de QF , a operação de quantização resulta em uma baixa qualidade de imagem e em maior compressão (dados com baixa taxa de *bits*), enquanto valores elevados apresentam um efeito contrário. Variou-se QF de 10 a 90 a passos de cinco unidades.

Para recuperar a imagem comprimida, o procedimento inverso é realizado. O bloco desquantizado é obtido multiplicando elemento por elemento as entradas do bloco quantizado pelos elementos da matriz de quantização. Subsequentemente, a transformada inversa dada em (2.28) foi computada. Para uma comparação quantitativa de desempenho, considerou-se como figuras de mérito a *relação sinal-ruído de pico* (PSNR) (BHASKARAN; KONSTANTINIDES, 1997; RAO; YIP, 2001; SALOMON; MOTTA, 2010) e a *similaridade estrutural* (SSIM) (WANG et al., 2004). O PSNR compara uma imagem original \mathbf{I} com a imagem recuperada \mathbf{R} de acordo com (SALOMON; MOTTA, 2010, p. 464)

$$\text{PSNR}(\mathbf{I}, \mathbf{R}) = 10 \cdot \log_{10} \left(\frac{L^2}{\text{MSE}_s(\mathbf{I}, \mathbf{R})} \right), \quad (2.30)$$

em que $L = 255$ é o valor máximo para a representação dos *pixels* (imagens representadas com

8 bits),

$$\text{MSE}_s(\mathbf{I}, \mathbf{R}) = \text{E} [(\mathbf{I} - \mathbf{R})^2] \quad (2.31)$$

é o erro médio quadrático de sinal ou energia do sinal de ruído (RAO; YIP, 2001) e $\text{E}[\cdot]$ retorna o valor médio.

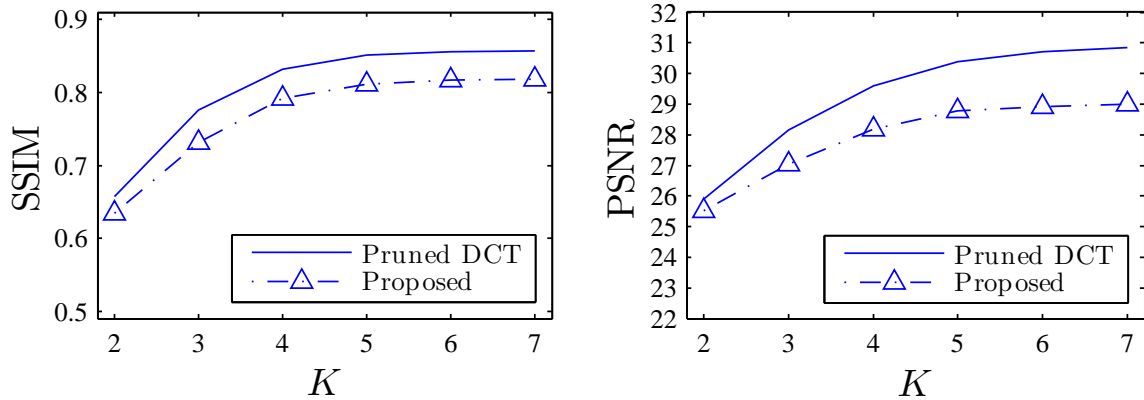
Entretanto, os valores de PSNR nem sempre correspondem com a qualidade de imagem percebida de fato pelo *sistema visual humano* (HVS)(BHASKARAN; KONSTANTINIDES, 1997, p. 9). O SSIM foi introduzido como uma métrica alternativa ao PSNR que tenta capturar melhor a qualidade visual percebida pelo HVS (WANG et al., 2004). O SSIM compara um bloco 8×8 da imagem original \mathbf{I} com sua respectiva versão recuperada \mathbf{R} de acordo com a seguinte expressão:

$$\text{SSIM}(\mathbf{I}, \mathbf{R}) = \frac{[2\mu_I\mu_R + (L \cdot 10^{-2})] \cdot [2\sigma_{IR} + (3L \cdot 10^{-2})]}{[\mu_I^2 + \mu_R^2 + (L \cdot 10^{-2})] \cdot [\sigma_I^2 + \sigma_R^2 + (3L \cdot 10^{-2})]}, \quad (2.32)$$

em que $\mu_I = \sum_{i=1}^8 \sum_{j=1}^8 \omega_{i,j} \cdot \mathbf{I}_{i,j}$, $\sigma_I = \left(\sum_{i=1}^8 \sum_{j=1}^8 \omega_{i,j} \cdot (\mathbf{I}_{i,j} - \mu_I)^2 \right)^{1/2}$, $\sigma_{IR} = \sum_{i=1}^8 \sum_{j=1}^8 \omega_{i,j} \cdot (\mathbf{I}_{i,j} - \mu_I) \cdot (\mathbf{R}_{i,j} - \mu_R)$ e $\omega_{i,j}$ é a entrada de uma função de ponderação gaussiana $\mathbf{w} = [\omega_{i,j}]$, $i, j = 1, 2, \dots, 8$, com desvio padrão de 1.5 e normalizada para soma unitária. A métrica SSIM é confinada ao intervalo $[0, 1]$, em que o valor 1 representa reconstrução perfeita, enquanto valores baixos indicam uma baixa qualidade de imagem.

O valor médio para cada métrica entre todas as imagens comprimidas foi computado. A Figura 3 mostra os resultados. Para $K \geq 4$, os métodos propostos apresentam um desempenho competitivo e comparável com os resultados da DTT exata. Para tal faixa de valores de K , a redução em desempenho é, em média e considerando a métrica SSIM, de 3.54% a uma redução média de custo de complexidade aditiva de 60.87%. Valores menores do parâmetro K são mais apropriados para altas taxas de compressão. Para $QF < 25$ e $K \in \{2, 3\}$, a redução em desempenho média é de 5.92%, enquanto a redução em complexidade aditiva média é de 80.54%. Na Figura 4, compara-se o desempenho dos métodos propostos com versões podadas da DCT exata (MAKKAOUI; LECUIRE; MOUREAUX, 2010; LECUIRE; MAKKAOUI; MOUREAUX, 2012) para cada valor de K e com $QF = 25$ — valores baixos de QF proporcionam alta compressão e $QF = 25$ ainda preserva uma satisfatória qualidade visual. A

Figura 4 – Métricas de desempenho em compressão de imagem tipo JPEG para os métodos podados propostos para a 2D DTT aproximada e métodos podados para a DCT exata para $QF = 25$.



Fonte: (COUTINHO et al., 2018)

redução média de SSIM dos métodos propostos é de 4.7% apenas. Uma comparação qualitativa é apresentada na Figura 5 e Figura 6 com $QF = 25$. As imagens são praticamente as mesmas.³

2.3.2 Compressão de Vídeo

As aproximações podadas propostas também tiveram o desempenho avaliado nos cenários de compressão de vídeo. Foram embutidas as 2D DTT exata, 2D DTT aproximada e suas versões podadas propostas nos codecs de código livre H.264/AVC (RICHARDSON, 2010) e H.265/HEVC (SULLIVAN et al., 2012a). Em ambos os casos, foram considerados para codificação onze vídeos no formato *common intermediate format* (CIF) e apresentando 300 quadros, obtidos da base de dados disponível em (XIPH...). Cada vídeo foi codificado empregando o software padrão e suas versões modificadas, em que as transformadas de comprimento oito foram substituídas pelas 2D DTT exata, 2D DTT aproximada e as versões podadas propostas. As matrizes diagonais associadas com cada transformada proposta são completamente absorvidas na etapa de quantização. Foram considerados valores fixos do parâmetro de quantização $QP = \{22, 27, 32, 37, 42\}$. Tal parâmetro ajusta o nível de quantização empregado. Quanto maior for o valor de QP , mais agressiva será a quantização e, portanto, uma maior compressão e uma menor taxa de *bits* são esperadas. Os valores médios de PSNR e SSIM para a componente de luminância e a taxa de bits, são computados pelo codec para os quadros reconstruídos.

³ Para uma melhor visualização das imagens, sugere-se a versão digitalizada deste trabalho.

Figura 5 – Avaliação qualitativa de desempenho para compressão de imagem ‘lena’ com $QF = 25$ para a 2D DCT e 2D DTT exatas, 2D DTT aproximada e suas versões podadas para cada parâmetro K .



(a) Original



(b) DCT exata (PSNR = 33.69, SSIM = 0.8871)



(c) DCT podada (PSNR = 32.98, SSIM = 0.8816)



(d) Exact DTT (PSNR = 33.33, SSIM = 0.8825)



(e) DTT exata podada (PSNR = 32.46, SSIM = 0.8752)



(f) DTT aproximada (PSNR = 31.14, SSIM = 0.8539)



(g) DTT aproximada podada ($K = 7$) (PSNR = 31.16, SSIM = 0.8540)



(h) DTT aproximada podada ($K = 6$) (PSNR = 31.16, SSIM = 0.8539)



(i) DTT aproximada podada ($K = 5$) (PSNR = 31.12, SSIM = 0.8532)



(j) DTT aproximada podada ($K = 4$) (PSNR = 30.59, SSIM = 0.8487)



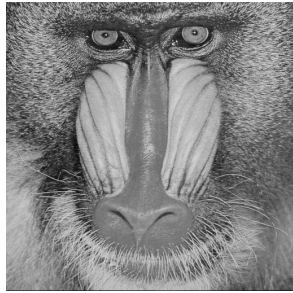
(k) DTT aproximada podada ($K = 3$) (PSNR = 29.34, SSIM = 0.8220)



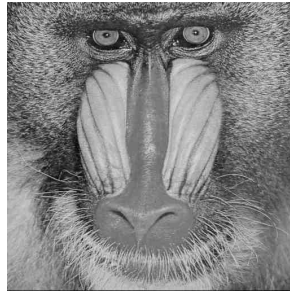
(l) DTT aproximada podada ($K = 2$) (PSNR = 27.44, SSIM = 0.7658)

Fonte: (COUTINHO et al., 2018)

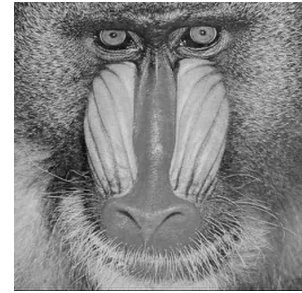
Figura 6 – Avaliação qualitativa de desempenho para compressão de imagem ‘mandrill’ com $QF = 25$ para a 2D DCT e 2D DTT exatas, 2D DTT aproximada e suas versões podadas para cada parâmetro K .



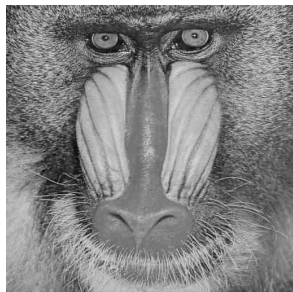
(a) Original



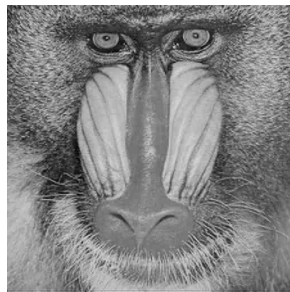
(b) DCT Exata (PSNR = 25.87, SSIM = 0.8111)



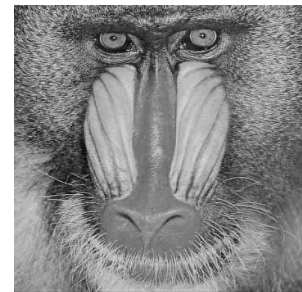
(c) DCT podada (PSNR = 23.72, SSIM = 0.7256)



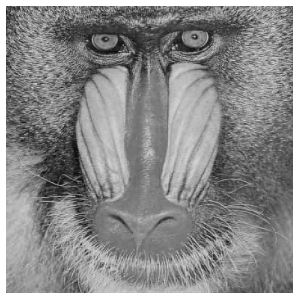
(d) DTT exata (PSNR = 25.54, SSIM = 0.7979)



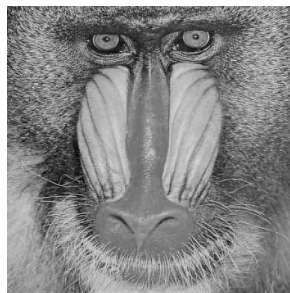
(e) DTT exata podada (PSNR = 23.47, SSIM = 0.7025)



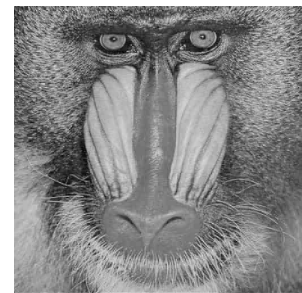
(f) DTT aproximada (PSNR = 24.33, SSIM = 0.7648)



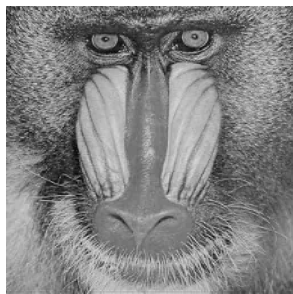
(g) DTT aproximada podada ($K = 7$) (PSNR = 24.33, SSIM = 0.7648)



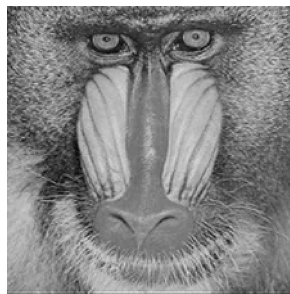
(h) DTT aproximada podada ($K = 6$) (PSNR = 24.19, SSIM = 0.7542)



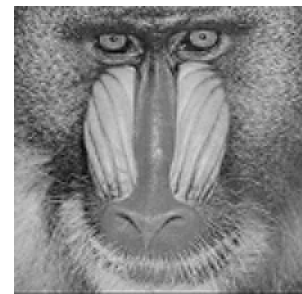
(i) DTT aproximada podada ($K = 5$) (PSNR = 23.84, SSIM = 0.7328)



(j) DTT aproximada podada ($K = 4$) (PSNR = 23.04, SSIM = 0.6830)

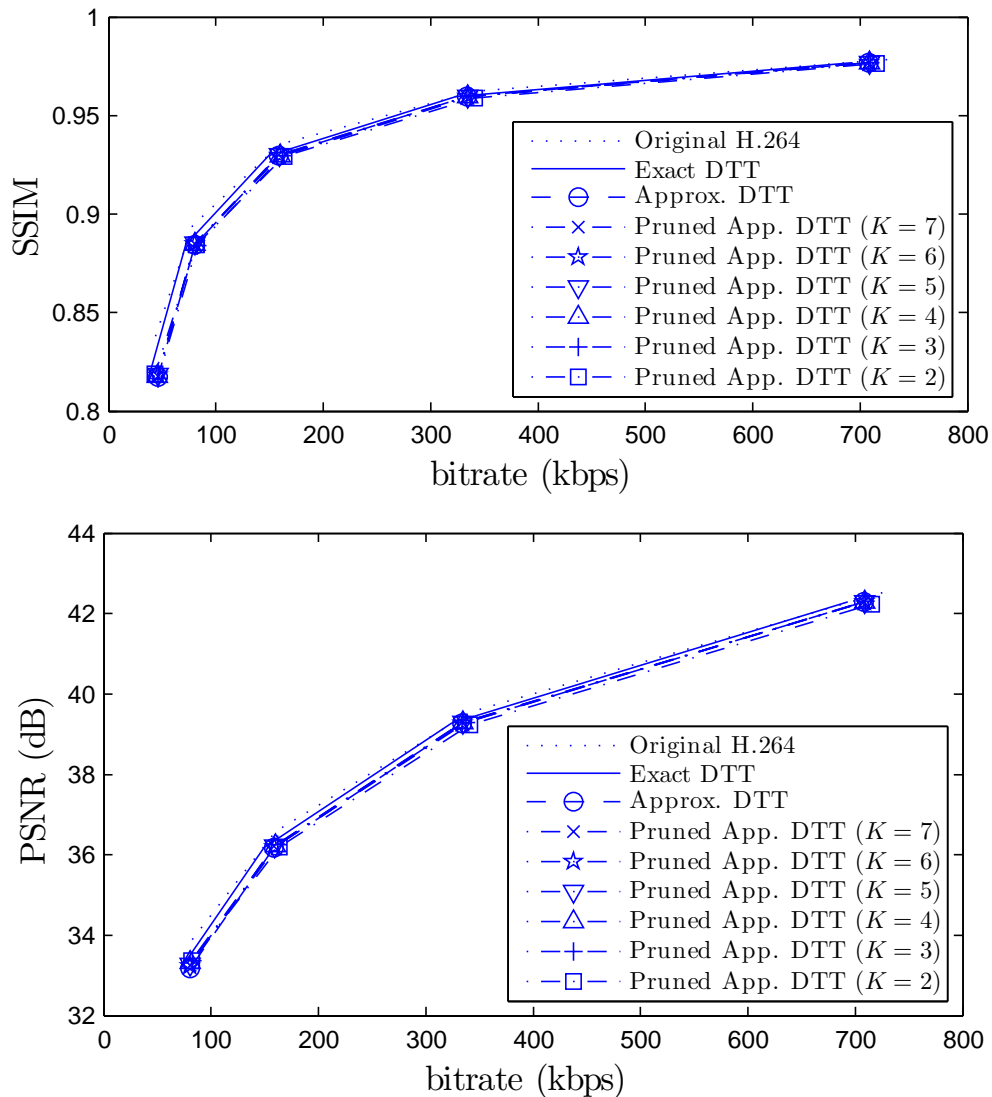


(k) DTT aproximada podada ($K = 3$) (PSNR = 21.88, SSIM = 0.5756)



(l) DTT aproximada podada ($K = 2$) (PSNR = 20.86, SSIM = 0.4455)

Figura 7 – Curvas de desempenho para H.264/AVC original e versões modificadas empregando as aproximações podadas para a 2D DTT para diferentes parâmetros de poda K .



Fonte: (COUTINHO et al., 2018)

2.3.2.1 H.264/AVC

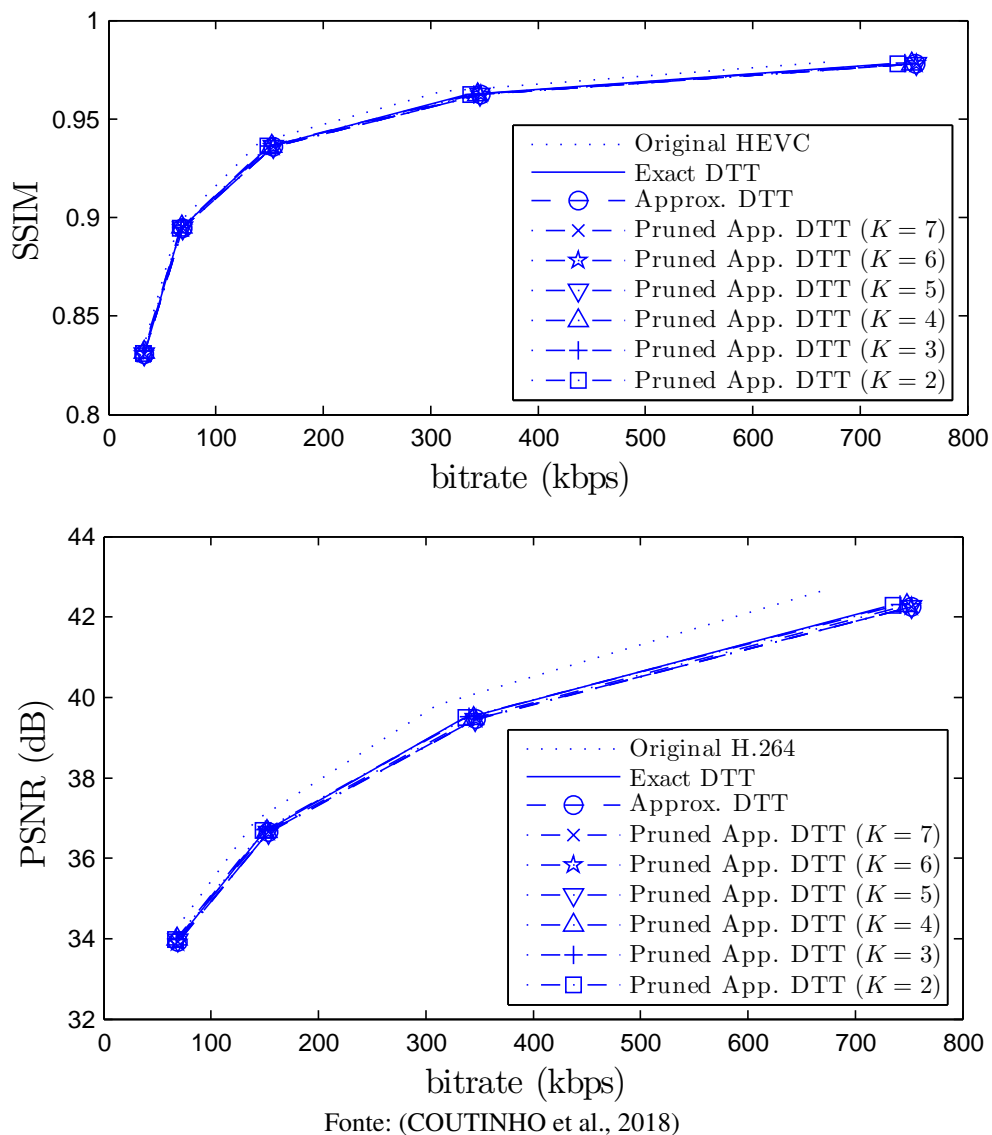
Para o codec H.264, foi empregada a biblioteca de software de código aberto x264 (TEAM, 2017) para codificar as sequências de vídeo consideradas. A transformada padrão presente no codec é uma DCT inteira (GORDON; MARPE; WIEGAND, 2004), cujo algoritmo rápido demanda 32 adições e 14 bit-shifts (GORDON; MARPE; WIEGAND, 2004). Sendo assim, o emprego das aproximações podadas propostas para a DTT leva a uma redução de 25% a 78% no custo aditivo no bloco de transformada de comprimento oito.

Os resultados quantitativos são mostrados na Figura 7 em termos de valores médios de

PSNR e SSIM para luminância entre todos os vídeos por taxa de bits obtidas para cada valor de QP . O emprego das transformadas podadas aproximadas propostas causa uma degradação desprezível na qualidade de vídeo obtida.

2.3.2.2 H.265/HEVC

Figura 8 – Curvas de desempenho para H.265/HEVC original e versões modificadas empregando as aproximações podadas para a 2D DTT para diferentes parâmetros de poda K .



As aproximações podadas propostas também foram avaliadas no contexto do codec HEVC/H.265 empregando a biblioteca de código aberto x265 disponível em (MULTICOREWARE, 2017). Em contraste com o H.264, tal codec emprega transformadas de comprimento 4, 8, 16

e 32. A transformada padrão de comprimento oito no HEVC demanda 22 multiplicações e 28 adições (BUDAGAVI et al., 2013), enquanto as aproximações propostas para a DTT requerem entre 14% e 75% menos operações de adição.

As figuras de mérito em função da taxa de bits para cada valor de QP são mostradas na Figura 8. Os vídeos codificados pelos codecs modificados com as aproximações podadas propostas foram praticamente idênticos aos vídeos codificados pelo codec original, a um custo computacional de transformada de comprimento oito consideravelmente inferior. De forma geral, os experimentos mostram que os codecs de vídeo considerados são robustos a modificações no bloco de transformada de comprimento oito apresentando tipicamente o mesmo desempenho relativa ao software padrão. Tal resiliência em relação a alterações no bloco de transformação 8×8 pode ser explicada pelo grande leque de técnicas empregadas em tais padrões de codificação de vídeo além dos blocos de transformada, como os algoritmos de estimação e compensação de movimento (SULLIVAN et al., 2012b). Dessa forma, o emprego de transformadas de muito baixa complexidade — tais como as aproximações propostas — é factível neste contexto prático.

2.4 ANÁLISE E DISTRIBUIÇÃO DE ENERGIA POR COEFICIENTE

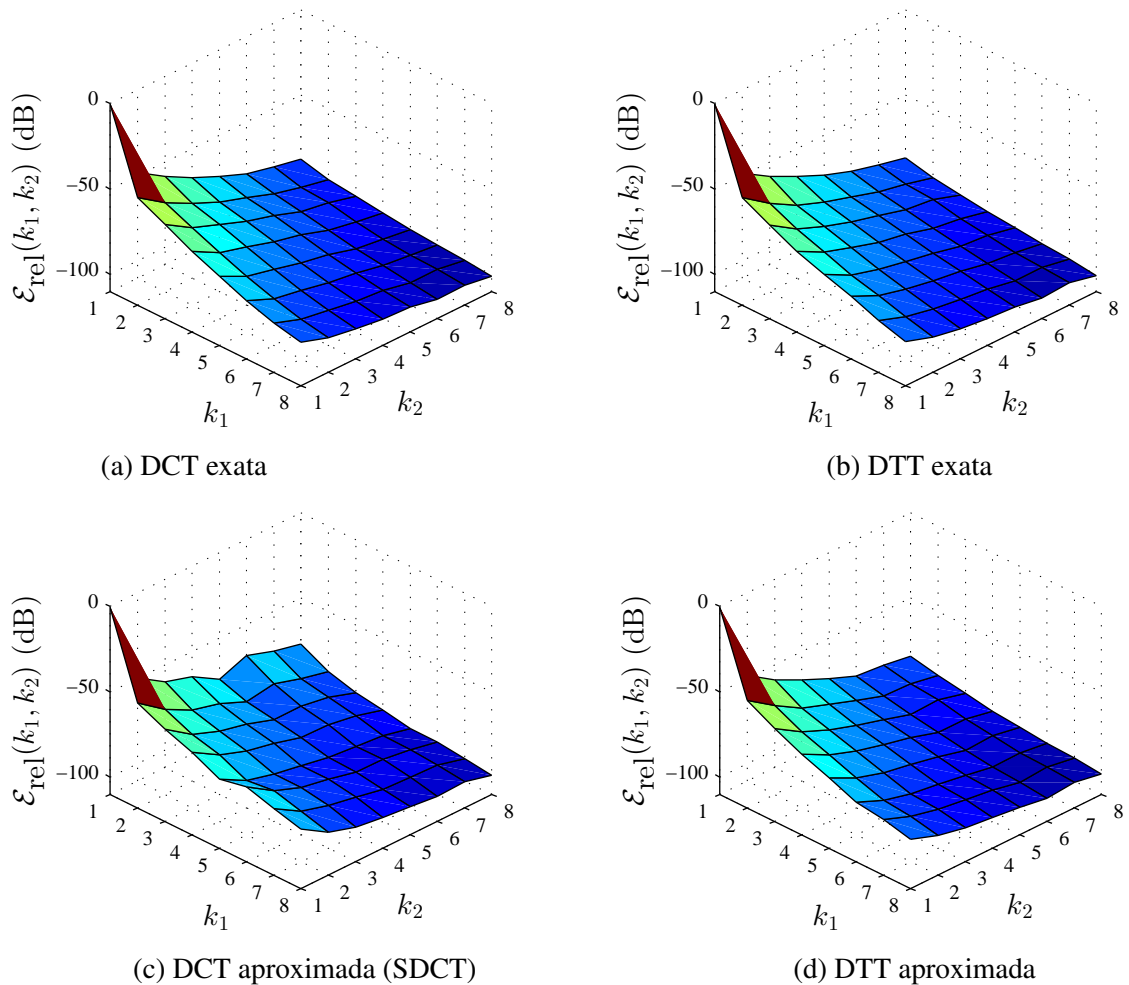
Para averiguar a capacidade de concentração de energia das aproximações podadas para a DTT propostas, foram empregadas as métricas de distribuição de energia média relativa por coeficiente, em dB, e a energia relativa média por zona, sugeridas em trabalho prévio (CONTINHO, 2015, p. 98). Seja \mathbf{Y} um bloco transformado de dimensão 8×8 — não podado — a energia relativa por coeficiente é dada por

$$\mathcal{E}_{\text{rel}}[k_1, k_2] = \frac{y[k_1, k_2]}{\mathcal{E}_Y},$$

em que $y[k_1, k_2]$ são as entradas de \mathbf{Y} e \mathcal{E}_Y é a energia total de \mathbf{Y} .

A energia relativa por zona representa $\mathcal{E}_{\text{zona}}[K]$ representa a razão entre a energia total concentrada pelo bloco \mathbf{Y}_K e a à energia total do bloco \mathbf{Y} não podado. Para comparação, também são analisadas as DCT exata, DTT exata, e a SDCT. São considerados os valores médios obtidos entre todos os blocos das 50 imagens consideradas do banco (Signal and Image Pro-

Figura 9 – Distribuição de energia por coeficiente para os métodos propostos para a 2D DTT e métodos usuais.

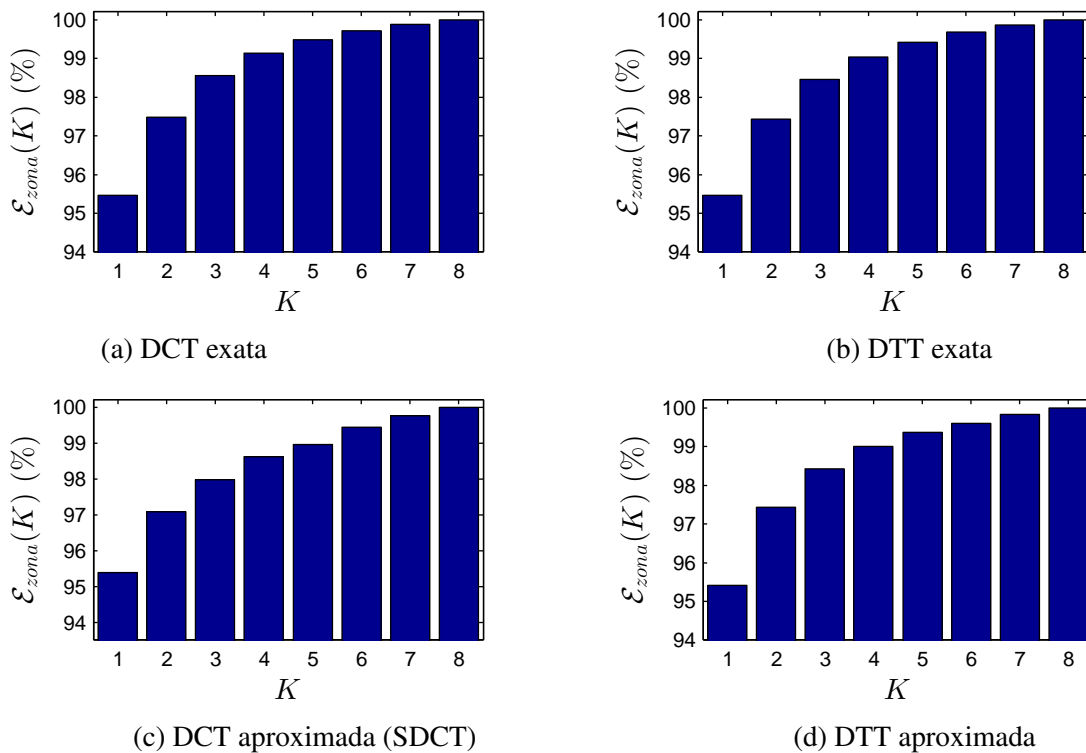


Fonte: o autor

cessing Institute, 2018). A Figura 9 mostra a energia relativa por coeficiente, em dB. A DTT aproximada mostra uma capacidade de concentração de energia similar aos casos da DCT exata e DTT exata.

A Figura 10 apresenta a distribuição de energia por zona. Ambas as DTT exata e DCT exata apresentam capacidade de concentrar mais de 99% de energia nos $K = 4$ coeficientes de baixa frequência, em média. A DTT aproximada também apresenta tal característica, enquanto a SDCT concentra mais de 99% de energia apenas para $K = 6$. Tal fato indica que os métodos propostos para a DTT apresentam capacidade de compactação de energia similares aos métodos exatos.

Figura 10 – Concentração de energia por zona para cada parâmetro de poda K para os métodos propostos e métodos usuais.



Fonte: o autor

2.5 IMPLEMENTAÇÃO EM FPGA

As aproximações podadas propostas para a DTT, $\hat{\mathbf{T}}_{(K)}$, $K = 2, \dots, 7$, assim como sua versão original não podada, $\hat{\mathbf{T}}$, foram implementadas em *field-programmable gate array* (FPGA) empregando a placa *Xilinx ML605 board* e o ambiente *MATLAB/Simulink*[®] integrado com a ferramenta *Xilinx System Generator for DSP*[®]. Foram consideradas oito entradas paralelas com dados de 8 bits—um comprimento típico de palavra para imagens (Signal and Image Processing Institute, 2018)—e um processamento empregando *pipeline*. A arquitetura proposta apresenta quatro estágios de adições e dois estágios de deslocamento de bits. Dessa forma, a latência do *pipeline* é de seis ciclos de relógio. As arquiteturas para os métodos de melhor desempenho em codificação ($K = 4, 5, 6$ e 7) são mostradas na Figura 11.

As arquiteturas desenvolvidas foram testadas empregando cem mil vetores de teste aleatórios de comprimento oito, gerados na plataforma MATLAB. Utilizando co-simulação em *hardware* via cabo JTAG, os vetores de testes foram enviados ao dispositivo FPGA, onde o

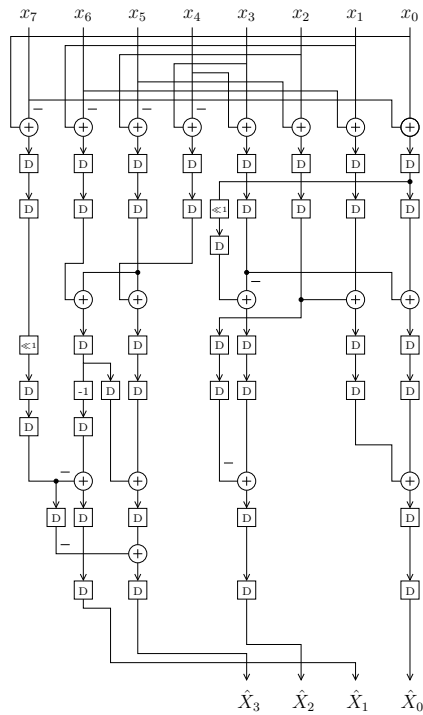
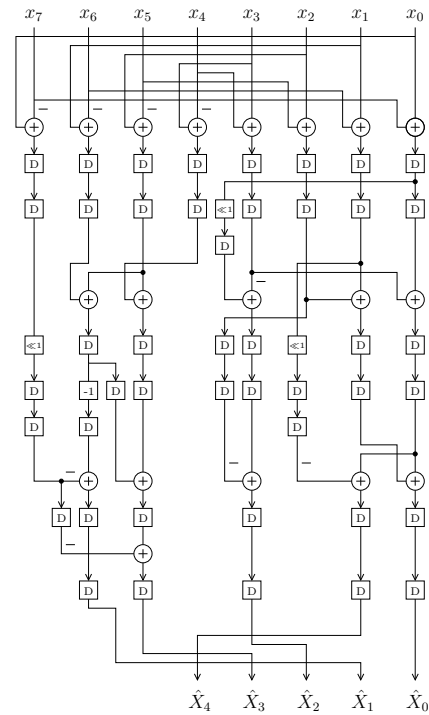
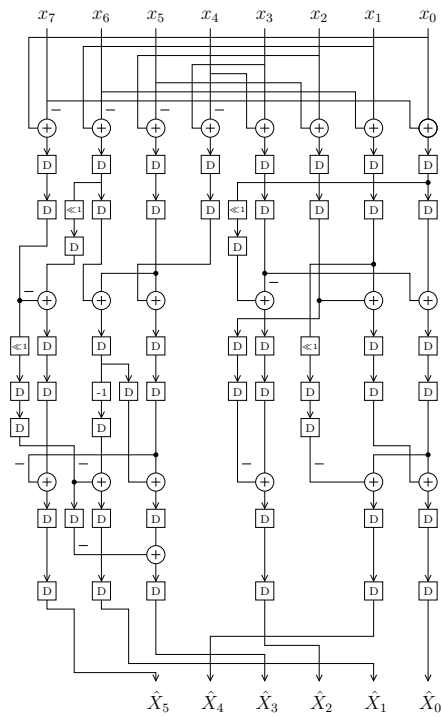
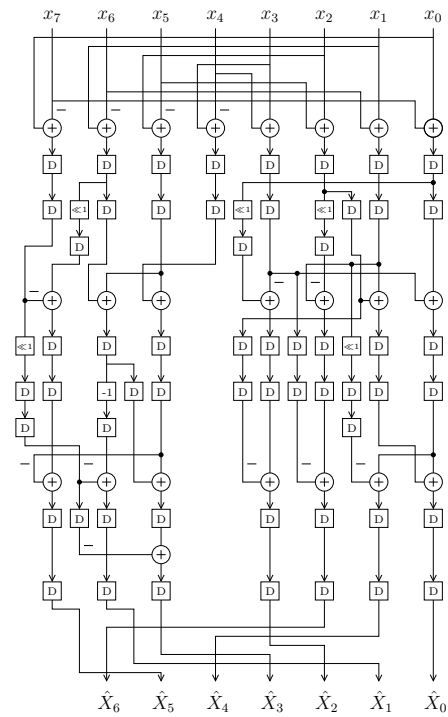
processamento foi realizado. Os vetores de saída obtidos pelo processamento em FPGA foram direcionados de volta para o ambiente MATLAB, onde foram verificados através de comparação com o algoritmo implementado em *software*.

Para avaliar quantitativamente as arquiteturas propostas, foram consideradas as seguintes métricas: número de *slices*, número de *flip-flops* (FF) e atraso de caminho crítico (T_{cpd}), em nanossegundos (ns). A máxima frequência de operação foi determinada através do atraso de caminho crítico de acordo com $F_{max} = (T_{cpd})^{-1}$, em megahertz (MHz). Os valores foram obtidos através das ferramentas de síntese e roteamento providas pela plataforma *Xilinx System Generator*. A potência estática P_e , em watts (W) e energia consumida por ciclo de relógio E_d , em nanojoule (nJ), foram estimadas através do *Xilinx XPower Analyzer*[®]. As figuras de mérito de *area-time* (AT) e *area-time-square* (AT^2) foram calculadas de modo que a métrica de área é medida em termos de número de *slices* e o tempo é dado pelo atraso de caminho crítico.

Os valores de todas as métricas mencionadas anteriormente são mostrados na Tabela 3. Como resultado de demandar um menor número de operações aritméticas, as aproximações podadas propostas para a DTT apresentam um menor consumo de *hardware*. Tal resultado é evidenciado pelas medidas AT e AT^2 , na Tabela 3. As reduções percentuais de tais métricas para cada método podado com fator de poda K em relação ao método original não podado são mostradas na Tabela 4 (o valor de referência é a DTT aproximada não podada ou $K = 8$).

Para comparação, foram implementadas diversas arquiteturas podadas para a DCT exata baseadas nos algoritmos podados propostos em (MAKKAOUI; LECUIRE; MOUREAUX, 2010; LECUIRE; MAKKAOUI; MOUREAUX, 2012), Tais algoritmos podados são versões podadas do algoritmo de Loeffler (LOEFFLER; LIGTENBERG; MOSCHYTZ, 1989). As métricas de *hardware* obtidas são mostradas na Tabela 5. É notável que o consumo de recursos de *hardware* é substancialmente maior para os métodos podados baseados na DCT exata quando comparados com os métodos podados propostos, mostrados na Tabela 3. As reduções percentuais de AT e AT^2 dos métodos podados propostos com fator de poda K comparados com as arquiteturas podadas baseadas na DCT podados também com fator de poda K são mostradas na Tabela 6 (o valor de referência é a DCT podada com fator K). Em média, as reduções de AT e AT^2 são de 81% e 83%, respectivamente.

Figura 11 – Arquiteturas realizadas para aproximações podadas para a 2D DTT com alguns fatores de poda K . O elemento D representa um *delay* ou registrador para sincronia do pipeline.

(a) $K = 4$ (b) $K = 5$ (c) $K = 6$ (d) $K = 7$

Fonte: (COUTINHO et al., 2018)

Tabela 3 – Consumo de recursos de *hardware* e potência para os métodos propostos para diferentes fatores de poda K

K	CLB	FF	T_{cpd} (ns)	F_{max} (MHz)	E_d (nJ)	P_e (W)	AT	AT^2
8	144	479	2.5	400	3.60	3.47	360	900
7	138	451	2.3	434	3.41	3.47	317	730
6	119	403	2.4	416	3.03	3.46	286	685
5	112	362	1.8	555	2.97	3.47	202	363
4	109	343	1.8	555	2.36	3.46	196	353
3	78	264	1.7	588	0.83	3.44	133	255
2	45	154	1.7	588	0.60	3.43	76	130

Fonte: (COUTINHO et al., 2018)

Tabela 4 – Reduções percentuais de AT e AT^2 de cada 2D DTT aproximada podada com fator de poda K em comparação com a DTT aproximada não podada ($K = 8$)

K	AT reduction	AT^2 reduction
7	11.94 %	18.89 %
6	20.56 %	23.89 %
5	43.89 %	59.67 %
4	45.56 %	60.78 %
3	63.05 %	71.67 %
2	78.89 %	85.56 %

Fonte: (COUTINHO et al., 2018)

Tabela 5 – Consumo de recursos de *hardware* e potência para a DCT exata podada para diferentes fatores de poda K

K	CLB	FF	T_{cpd} (ns)	F_{max} (MHz)	E_d (nJ)	P_e (W)	AT	AT^2
8	753	3349	3.6	278	0.68	3.44	2711	9758
7	649	2762	2.8	357	0.61	3.44	1817	5088
6	547	2440	2.8	357	1.17	8.30	1532	4288
5	441	1860	2.3	435	0.92	3.45	1014	2333
4	467	1830	2.2	455	1.29	3.46	1027	2260
3	314	1249	2.1	476	0.87	3.44	659	1385
2	217	879	1.9	576	0.33	5.83	412	783

Fonte: (COUTINHO et al., 2018)

Tabela 6 – Reduções percentuais de AT e AT^2 de cada aproximação para a 2D DTT com fator de poda K em comparação com a DCT exata podada também com fator de poda K

K	AT reduction	AT^2 reduction
7	82.55 %	85.65 %
6	81.33 %	84.02 %
5	80.08 %	84.44 %
4	80.91 %	84.38 %
3	79.82 %	81.59 %
2	81.55 %	83.40 %

Fonte: (COUTINHO et al., 2018)

3 APROXIMAÇÕES PARA A DCT MULTIDIMENSIONAL

Vídeos digitais consistem em uma sequência de quadros de imagens estáticas, isto é, uma sequência de arranjos bidimensionais de *pixels* (NATARAJAN; AHMED, 1977). Consequentemente, um vídeo digital pode ser interpretado como um sinal digital tridimensional, ou um arranjo volumétrico de dados, com duas dimensões espaciais e uma dimensão temporal (CHAN; LEE, 1997).

Cada quadro de vídeo é amostrado em intervalos de tempo frequentes o suficiente para que o sistema visual humano não perceba distorções (NATARAJAN; AHMED, 1977). Por exemplo, vídeos em resolução 4K são considerados usualmente uma taxa de quadros de 30 quadros por segundo, ou 30 fps (FAN et al., 2018). Consequentemente, os *pixels* quadros vizinhos tendem a apresentar uma alta correlação (GHARAVI; MILLS, 1990). Em outras palavras, o sinal digital tridimensional apresenta correlação espacial, propriedade de imagens estáticas naturais e com significado ao olho humano (RAO; YIP, 2001; WALLACE, 1992).

Os padrões de codificação de vídeo dominantes, tais como MPEG, H.264 e HEVC, consideram transformadas bidimensionais para a redução de redundâncias espaciais, isto é, a correlação entre *pixels* de um mesmo quadro. Técnicas de codificação deste tipo são denominadas de *codificação intraquadros* (DUFAUX; MOSCHENI, 1995). Neste tipo de codificador, as redundâncias temporais são tratadas através de algoritmos de *compensação de movimento* e *estimação de movimento* (GHARAVI; MILLS, 1990; DUFAUX; KONRAD, 2000). Tais algoritmos pertencem a classe de técnicas de codificação preditiva (DUFAUX; MOSCHENI, 1995), que apresentam dependências entre os quadros. Dessa forma erros no fluxo de *bits* podem levar a perda de sincronismo e de qualidade. Além disso, o custo computacional envolvido em tais algoritmos dificulta a implementação de codificadores em tempo real, especialmente para contextos com restrição acentuada de recursos (SAWANT; ADJEROH, 2011).

Alternativas para codificação de vídeo que evitem tais algoritmos complexos são, portanto, desejáveis. Uma abordagem que pode ser empregada é codificação tridimensional de vídeo, também chamada de codificação *entre quadros* (NATARAJAN; AHMED, 1977; SAWANT;

ADJEROH, 2011; DUFAUX; MOSCHENI, 1995). Este método se baseia em uma transformada tridimensional com capacidade de descorrelação espaço-temporal, Diversos esquemas baseados em codificação entre quadros foram propostos, evitando algoritmos de estimação de movimento, codificação preditiva e propagação de erro (CHAN; SIU, 1997). Além disso, codificadores tridimensionais apresentam a vantagem de serem simétricos, isto é, o codificador e o decodificador uma apresentam estrutura similar (CHAN; LEE, 1997): transformada, quantização e codificação; decodificação, desquantização e transformada inversa.

Entretanto, codificadores entre quadros apresentam a desvantagem de apresentar maior atraso e demandar mais memória (NATARAJAN; AHMED, 1977; CHAN; SIU, 1997). Sendo assim, algoritmos de baixa complexidade e baseados em representação inteira são vantajosos em tal contexto. Aproximações para a DCT livres de multiplicação e com baixíssimo custo computacional vêm sendo propostas nos últimos anos (HAWHEEL, 2001; LENGWEHASATIT; ORTEGA, 2004; CINTRA; BAYER, 2011; BAYER; CINTRA, 2012; BOUGUEZEL; AHMAD; SWAMY, 2008; BOUGUEZEL; AHMAD; SWAMY, 2009; BOUGUEZEL; AHMAD; SWAMY, 2013; POTLURI et al., 2014; CINTRA, 2011; CINTRA; BAYER; TABLADA, 2014; MADANAYAKE et al., 2015). Entretanto, ao nosso melhor conhecimento, aproximações para a DCT de mais altas dimensões representam um campo de pesquisa inexplorado.

3.1 REVISÃO MATEMÁTICA

Nesta Seção, são revisados conceitos matemáticos necessários relacionados à DCT e a teoria de tensores.

3.1.1 1D e 2D DCT

A DCT mapeia um sinal discreto $\mathbf{x} = [x[0] \ x[1] \ \dots \ x[N-1]]^\top$ de comprimento N em um sinal discreto $\mathbf{y} = [y[0] \ y[1] \ \dots \ y[N-1]]^\top$ de acordo com a seguinte relação (AHMED; NATARAJAN; RAO, 1974):

$$y[k] \triangleq \alpha_N[k] \cdot \sum_{n=0}^{N-1} x[n] \cdot \cos\left(\frac{\pi(2n+1)k}{2N}\right), \quad (3.1)$$

$$k = 0, 1, \dots, N-1,$$

em que

$$\alpha_N[k] \triangleq \sqrt{\frac{1}{N}} \cdot \begin{cases} 1, & \text{se } k = 0, \\ \sqrt{2}, & \text{caso contrário.} \end{cases}$$

Seja \mathbf{X} um sinal 2D de dimensões $N_1 \times N_2$, cujas entradas são dadas por $x[n_1, n_2]$, para $n_i = 0, 1, \dots, N_i - 1$, e $i = 1, 2$. As entradas do sinal bidimensional no domínio da transformada da 2D DCT são computadas de acordo com a seguinte expressão (CHO, 1991):

$$y[k_1, k_2] \triangleq \alpha_{N_1}[k_1] \cdot \alpha_{N_2}[k_2] \cdot \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x[n_1, n_2] \cdot \cos\left(\frac{\pi(2n_1 + 1)k_1}{2N_1}\right) \cdot \cos\left(\frac{\pi(2n_2 + 1)k_2}{2N_2}\right), \quad (3.2)$$

$$k_i = 0, 1, \dots, N_i - 1, \quad i = 1, 2.$$

Ambas a DCT e a 2D DCT podem ser expressas em termos de produtos de matrizes. Para o caso 1D, tem-se que

$$\mathbf{y} = \mathbf{C}_N \cdot \mathbf{x}, \quad (3.3)$$

em que \mathbf{C}_N é a matriz de transformação da DCT cujas entradas são dadas por

$$c_N[k, n] = \alpha_N[k] \cdot \cos\left(\frac{\pi(2n + 1)k}{2N}\right),$$

$$k, n = 0, 1, \dots, N - 1.$$

Para o caso bidimensional, o sinal 2D de entrada pode ser interpretado como uma matriz \mathbf{X} de dimensão $N_1 \times N_2$ e o sinal 2D transformado associado é dado por:

$$\mathbf{Y} = \mathbf{C}_{N_1} \cdot \mathbf{X} \cdot \mathbf{C}_{N_2}^\top. \quad (3.4)$$

3.1.2 3D DCT e Tensores de Alta Ordem

A 3D DCT de um sinal discreto \mathcal{X} com entradas $x[n_1, n_2, n_3]$, $n_i = 0, 1, \dots, N_i - 1$, para $i = 1, 2, 3$, é dada pelo sinal \mathcal{Y} , cujas entradas são dadas por (BOUSSAKTA, 2004; LI et

al., 2013):

$$\begin{aligned}
y[k_1, k_2, k_3] &\triangleq \alpha_{N_1}[k_1] \cdot \alpha_{N_2}[k_2] \cdot \alpha_{N_3}[k_3] \cdot \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \sum_{n_3=0}^{N_3-1} x[n_1, n_2, n_3] \cdot \cos\left(\frac{\pi(2n_1+1)k_1}{2N_1}\right) \\
&\quad \cdot \cos\left(\frac{\pi(2n_2+1)k_2}{2N_2}\right) \cdot \cos\left(\frac{\pi(2n_3+1)k_3}{2N_3}\right), \\
k_i &= 0, 1, \dots, N_i - 1, \quad i = 1, 2, 3.
\end{aligned} \tag{3.5}$$

A 2D e a 3D DCT representam transformadas separáveis com núcleo regular, de acordo com a Definição 2.3, com $\phi[n, k] = \left(\frac{\pi(2n+1)k}{2N}\right)$.

Vetores e matrizes podem ser modelados como tensores de primeira e segunda ordem, respectivamente (LATHAUWER; MOOR; VANDEWALLE, 2000). Analogamente, um sinal tridimensional pode ser interpretado como um tensor de terceira ordem (LATHAUWER; MOOR, 1998; NORTHCOTT, 2008; LI et al., 2013). Um tensor de ordem R -ésima pode ser interpretado como um arranjo cujas entradas são indexadas por R índices. De forma mais formal, um tensor pode ser definido como mapeamento sobre um conjunto de espaços vetoriais (LATHAUWER; MOOR, 1998).

Definição 3.1. *Seja $\mathcal{X} \in \mathbb{F}^{N_1 \times N_2 \times \dots \times N_R}$ um tensor de ordem R -ésima cujas entradas são dadas por $x[n_1, n_2, \dots, n_R]$, em que \mathbb{F} pode ser tanto o conjunto dos números reais ou complexos, e $n_i = 0, 1, \dots, N_i - 1$, para $i = 1, 2, \dots, R$. O produto em modo i -ésimo do tensor \mathcal{X} por uma matriz $\mathbf{M} \in \mathbb{F}^{H \times N_i}$ (BERNSTEIN, 2009, p. xxxv), denotado por $\mathcal{X} \times_i \mathbf{M}$, é definido como o tensor $\mathcal{Y} \in \mathbb{F}^{N_1 \times N_2 \times \dots \times N_{i-1} \times H \times N_{i+1} \times \dots \times N_R}$, cujas entradas são expressas pela seguinte expressão:*

$$y[n_1, \dots, n_{i-1}, h, n_{i+1}, \dots, n_R] \triangleq \sum_{n_i=0}^{N_i-1} x[n_1, \dots, n_i, \dots, n_R] \cdot m[h, n_i], \tag{3.6}$$

em que $m[h, n_i]$ são as entradas de \mathbf{M} e $h = 0, 1, \dots, H - 1$.

O produto em i -ésimo modo dado em (3.6) generaliza produtos de vetores e matrizes. De fato, considerando um vetor-coluna (tensor de primeira ordem) $\mathbf{v} \in \mathbb{F}^{N_1}$ e matrizes (tensores de segunda ordem) $\mathbf{A} \in \mathbb{F}^{N_1 \times N_2}$, $\mathbf{M}_1 \in \mathbb{F}^{L \times N_1}$, e $\mathbf{M}_2 \in \mathbb{F}^{H \times N_2}$, as seguintes representações

são verdadeiras:

$$\begin{aligned}\mathbf{v} \times_1 \mathbf{M}_1 &= \mathbf{M}_1 \cdot \mathbf{v}, \\ \mathbf{A} \times_1 \mathbf{M}_1 &= \mathbf{M}_1 \cdot \mathbf{A}, \\ \mathbf{A} \times_2 \mathbf{M}_2 &= \mathbf{A} \cdot \mathbf{M}_2^\top.\end{aligned}$$

Proposição 3.1. *Sejam as matrizes $\mathbf{M}_1 \in \mathbb{F}^{L \times N_i}$, $\mathbf{M}_2 \in \mathbb{F}^{H \times N_j}$ e o tensor $\mathcal{X} \in \mathbb{F}^{N_1 \times N_2 \times \dots \times N_R}$. Então, se $i \neq j$, é válida a seguinte propriedade (LATHAUWER; MOOR; VANDEWALLE, 2000):*

$$(\mathcal{X} \times_i \mathbf{M}_1) \times_j \mathbf{M}_2 = (\mathcal{X} \times_j \mathbf{M}_2) \times_i \mathbf{M}_1.$$

Demonstração. Sejam os tensores

$$\begin{aligned}\mathcal{X}_1 &\triangleq \mathcal{X} \times_i \mathbf{M}_1, \\ \mathcal{X}_2 &\triangleq \mathcal{X} \times_j \mathbf{M}_2, \\ \mathcal{Y} &\triangleq \mathcal{X} \times_i \mathbf{M}_1 \times_j \mathbf{M}_2.\end{aligned}\tag{3.7}$$

As entradas do tensor \mathcal{X}_1 são dadas por

$$x_1[n_1, \dots, l, \dots, n_R] = \sum_{n_i} x[n_1, \dots, n_i, \dots, n_R] \cdot m_1[l, n_i],\tag{3.8}$$

em que $m_1[l, n_i]$ e $x[n_1, \dots, n_R]$ são as entradas de \mathbf{M}_1 e \mathcal{X} , respectivamente. Note que $\mathcal{Y} = \mathcal{X}_1 \times_j \mathbf{M}_2$. Então as entradas de \mathcal{Y} são dadas por

$$y[n_1, \dots, l, \dots, h, \dots, n_R] = \sum_{n_j} x_1[n_1, \dots, l, \dots, n_j, \dots, n_R] \cdot m_2[h, n_j],\tag{3.9}$$

em que $m_2[h, n_j]$ são as entradas de \mathbf{M}_2 . Substituindo (3.8) em (3.9), tem-se

$$\begin{aligned}y[n_1, \dots, l, \dots, h, \dots, n_R] &= \sum_{n_j} \left(\sum_{n_i} x[n_1, \dots, n_i, \dots, n_j, \dots, n_R] \cdot m_1[l, n_i] \right) \cdot m_2[h, n_j] \\ &= \sum_{n_i} \left(\sum_{n_j} x[n_1, \dots, n_i, \dots, n_j, \dots, n_R] \cdot m_2[h, n_j] \right) \cdot m_1[l, n_i]\end{aligned}\tag{3.10}$$

$$= \sum_{n_i} x_2[n_1, \dots, h, \dots, n_i, \dots, n_R] \cdot m_1[l, n_i],\tag{3.11}$$

em que (3.10) é obtida invertendo as ordens dos somatórios e observando que

$$x_2[n_1, \dots, h, \dots, n_R] = \sum_{n_j} x[n_1, \dots, n_j, \dots, n_R] \cdot m_2[h, n_j] \quad (3.12)$$

são as entradas do tensor \mathcal{X}_2 . Note que (3.11) também representa as entradas do tensor $\mathcal{X}_2 \times_i \mathbf{M}_1$, portanto

$$\begin{aligned} \mathcal{Y} &= \mathcal{X}_2 \times_i \mathbf{M}_1 \\ &= \mathcal{X} \times_j \mathbf{M}_2 \times_i \mathbf{M}_1. \end{aligned} \quad (3.13)$$

Igualando (3.7) e (3.13), conclui-se a demonstração. □

Proposição 3.2. *Considere duas matrizes $\mathbf{M} \in \mathbb{F}^{H \times L}$ e $\mathbf{G} \in \mathbb{F}^{L \times N_i}$, e um tensor de ordem R -ésima $\mathcal{X} \in \mathbb{F}^{N_1 \times N_2 \times \dots \times N_i \times \dots \times N_R}$. Então, é válida a seguinte relação (LATHAUWER; MOOR; VANDEWALLE, 2000):*

$$\mathcal{X} \times_i (\mathbf{M} \cdot \mathbf{G}) = \mathcal{X} \times_i \mathbf{G} \times_i \mathbf{M}.$$

Demonstração. Sejam

$$\mathbf{P} \triangleq \mathbf{M} \cdot \mathbf{G},$$

$$\mathcal{X}_1 \triangleq \mathcal{X} \times_i \mathbf{G},$$

$$\mathcal{Y} \triangleq \mathcal{X} \times_i \mathbf{G} \times_i \mathbf{M}.$$

As entradas da matriz \mathbf{P} são dadas de acordo com o produto matricial usual, conforme:

$$p[h, n_i] = \sum_l m[h, l] \cdot g[l, n_i], \quad (3.14)$$

em que $m[h, l]$ e $g[l, n_i]$ são as entradas de \mathbf{M} e \mathbf{G} , respectivamente. Por sua vez, as entradas de \mathcal{X}_1 são dadas por

$$x_1[n_1, \dots, l, \dots, n_R] = \sum_{n_i} x[n_1, \dots, n_i, \dots, n_R] \cdot g[l, n_i], \quad (3.15)$$

em que $x[n_1, \dots, n_R]$ são as entradas de \mathcal{X} . Note que $\mathcal{Y} = \mathcal{X}_1 \times_i \mathbf{M}$. Logo, as entradas de \mathcal{Y} são dadas por

$$y[n_1, \dots, h, \dots, n_R] = \sum_l x_1[n_1, \dots, l, \dots, n_R] \cdot m[h, l]. \quad (3.16)$$

Substituindo (3.15) em (3.16), tem-se

$$y[n_1, \dots, h, \dots, n_R] = \sum_l \left(\sum_{n_i} x[n_1, \dots, n_i, \dots, n_R] \cdot g[l, n_i] \right) \cdot m[h, l]. \quad (3.17)$$

Invertendo as ordens dos somatórios em (3.17), reorganizando os termos e levando em conta (3.14), tem-se

$$\begin{aligned} y[n_1, \dots, h, \dots, n_R] &= \sum_{n_i} x[n_1, \dots, n_i, \dots, n_R] \left(\sum_l m[h, l] \cdot g[l, n_i] \right) \\ &= \sum_{n_i} x[n_1, \dots, n_i, \dots, n_R] \cdot p[h, n_i]. \end{aligned} \quad (3.18)$$

Note que (3.18) expressa as entradas do tensor $\mathcal{X} \times_i \mathbf{P}$. Portanto, tem-se que

$$\begin{aligned} \mathcal{X} \times_i \mathbf{P} &= \mathcal{Y} \\ \mathcal{X} \times_i (\mathbf{M} \cdot \mathbf{G}) &= \mathcal{X} \times_i \mathbf{G} \times_i \mathbf{M}, \end{aligned}$$

concluindo a demonstração. □

Em particular, é verdade que:

$$\mathcal{T} \times_i \mathbf{I}_{N_i} = \mathcal{T}. \quad (3.19)$$

Tendo em vista os conceitos introduzidos previamente, a 3D DCT pode ser expressa de acordo com produtos em modo i -ésimo por matrizes de transformação da DCT. Para o caso unidimensional, a Eq. (3.3) pode se escrita como $\mathbf{y} = \mathbf{x} \times_1 \mathbf{C}_N$. De forma semelhante, a 2D DCT em (3.4) é dada por $\mathbf{Y} = \mathbf{X} \times_1 \mathbf{C}_{N_1} \times_2 \mathbf{C}_{N_2}$. Sendo assim, seja $\mathcal{X} \in \mathbb{F}^{N_1 \times N_2 \times N_3}$ o sinal 3D discreto de entrada considerado como um tensor de terceira ordem. A 3D DCT é dada pelo também tensor de terceira ordem de acordo com (LI et al., 2013)

$$\mathcal{Y} = \mathcal{X} \times_1 \mathbf{C}_{N_1} \times_2 \mathbf{C}_{N_2} \times_3 \mathbf{C}_{N_3}.$$

3.1.3 Aproximações para a DCT

O principal objetivo das DCTs aproximadas é obter propriedades matemáticas similares a DCT exata a um custo computacional significativamente menor. Usualmente, uma aproximação para a DCT de comprimento N é representada por uma matriz de transformação $\hat{\mathbf{C}}_N$,

que em geral resulta de um produto de uma matriz de baixa complexidade \mathbf{T}_N e uma matriz diagonal \mathbf{S}_N .

Propriedades de ortonormalidade ou quase-ortonormalidade são asseguradas pela matriz diagonal $\mathbf{S}_N = \sqrt{[\text{diag}(\mathbf{T}_N \cdot \mathbf{T}_N^\top)]^{-1}}$, conforme discutido na Seção 2.1.4. Sendo assim, $\hat{\mathbf{C}}_N = \mathbf{S}_N \cdot \mathbf{T}_N$ (HAWHEEL, 2001; LENGWEHASATIT; ORTEGA, 2004; CINTRA; BAYER, 2011; BAYER; CINTRA, 2012; BOUGUEZEL; AHMAD; SWAMY, 2008; BOUGUEZEL; AHMAD; SWAMY, 2009; BOUGUEZEL; AHMAD; SWAMY, 2013; POTLURI et al., 2014). Considerando o caso unidimensional, um vetor de entrada \mathbf{x} de comprimento N é transformado no vetor de saída \mathbf{y} também de comprimento N de acordo com a seguinte expressão:

$$\begin{aligned} \mathbf{y} &= \hat{\mathbf{C}}_N \cdot \mathbf{x} \\ &= \mathbf{S}_N \cdot \mathbf{T}_N \cdot \mathbf{x}, \end{aligned} \tag{3.20}$$

em que tais matrizes são quadradas de ordem N . A transformação inversa é calculada de acordo com $\mathbf{x} = \hat{\mathbf{C}}_N^{-1} \cdot \mathbf{y}$, em que a matriz inversa é dada por (CINTRA; BAYER; TABLADA, 2014)

$$\hat{\mathbf{C}}_N^{-1} = \begin{cases} \mathbf{T}_N^\top \cdot \mathbf{S}_N, & \text{se } \hat{\mathbf{C}}_N \text{ é ortonormal,} \\ \mathbf{T}_N^{-1} \cdot \mathbf{S}_N^{-1}, & \text{caso contrário.} \end{cases} \tag{3.21}$$

Para 2D DCTs aproximadas, uma matriz de entrada \mathbf{X} de dimensões $N \times N$ é submetida à seguinte transformação

$$\begin{aligned} \mathbf{Y} &= \hat{\mathbf{C}}_N \cdot \mathbf{X} \cdot \hat{\mathbf{C}}_N^\top \\ &= \mathbf{S}_N \cdot \mathbf{T}_N \cdot \mathbf{X} \cdot \mathbf{T}_N^\top \cdot \mathbf{S}_N^\top \\ &= (\mathbf{s}_N \cdot \mathbf{s}_N^\top) \odot (\mathbf{T}_N \cdot \mathbf{X} \cdot \mathbf{T}_N^\top), \end{aligned}$$

em que \mathbf{s}_N é um vetor coluna de comprimento N que contém os elementos da diagonal da matriz \mathbf{S}_N , \mathbf{Y} é a matriz de saída de dimensões $N \times N$ contendo os coeficientes no domínio da transformada, e \odot denota o produto de Hadamard (HORN; JOHNSON, 2012). O termo $\mathbf{s}_N \cdot \mathbf{s}_N^\top$ é uma matriz repleta de elementos multiplicativos, em que a operação $\mathbf{T}_N \cdot \mathbf{X} \cdot \mathbf{T}_N^\top$ é frequentemente livre de multiplicação ou com muito baixo custo computacional.

Em alguns contextos, o termo $\mathbf{s}_N \cdot \mathbf{s}_N^\top$ pode ser embutido em uma etapa subsequente de um sistema maior. Por exemplo, em se tratando de compressão de imagem e vídeo, a etapa de

Tabela 7 – Aproximações para a DCT de comprimento $N = 8$ encontradas na literatura, com matriz de transformação dada por $\hat{\mathbf{C}}_8 = \mathbf{S}_8 \cdot \mathbf{T}_8$

Método	\mathbf{T}_8	\mathbf{S}_8
SDCT	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}$	$\text{diag} \left(\frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}} \right).$
LODCT	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & -1 & -1 & -1 \\ 1 & \frac{1}{2} & -\frac{1}{2} & -1 & -1 & -\frac{1}{2} & \frac{1}{2} & 1 \\ 1 & 0 & -1 & -1 & 1 & 1 & 0 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 \\ \frac{1}{2} & -1 & 1 & -\frac{1}{2} & -\frac{1}{2} & 1 & -1 & \frac{1}{2} \\ 0 & -1 & 1 & -1 & 1 & -1 & 1 & 0 \end{bmatrix}$	$\text{diag} \left(\frac{1}{\sqrt{8}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{6}} \right)$
RDCT	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & -1 & -1 & 0 & 0 & 1 \\ 1 & 0 & -1 & -1 & 1 & 1 & 0 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 \\ 0 & -1 & 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & -1 & 1 & -1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$	$\text{diag} \left(\frac{1}{\sqrt{8}}, \frac{1}{\sqrt{6}}, \frac{1}{2}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{6}}, \frac{1}{2}, \frac{1}{\sqrt{6}} \right)$
MRDCT	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \end{bmatrix}$	$\text{diag} \left(\frac{1}{\sqrt{8}}, \frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{2}} \right)$
BAS-2008	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 \\ 1 & \frac{1}{2} & -\frac{1}{2} & -1 & -1 & -\frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 1 & -1 \\ \frac{1}{2} & -1 & 1 & -\frac{1}{2} & -\frac{1}{2} & 1 & -1 & \frac{1}{2} \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \end{bmatrix}$	$\text{diag} \left(\frac{1}{\sqrt{8}}, \frac{1}{2}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{8}}, \frac{1}{2}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{2}} \right)$
BAS-2009	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \end{bmatrix}$	$\text{diag} \left(\frac{1}{\sqrt{8}}, \frac{1}{2}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{8}}, \frac{1}{2}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{2}} \right)$
BAS-2013	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}$	$\text{diag} \left(\frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}} \right)$
IADCT	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 & -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}$	$\text{diag} \left(\frac{1}{\sqrt{8}}, \frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{2}} \right)$

Fonte: o autor

quantização pode incorporar completamente tais elementos multiplicativos de $\mathbf{s}_N \cdot \mathbf{s}_N^\top$ (BRITANAK; YIP; RAO, 2007). Neste caso, o termo $\mathbf{s}_N \cdot \mathbf{s}_N^\top$ não introduz nenhuma operação aritmética no sistema completo (HAWHEEL, 2001; LENGWEHASATIT; ORTEGA, 2004; CINTRA; BAYER, 2011; BAYER; CINTRA, 2012; BOUGUEZEL; AHMAD; SWAMY, 2008; BOUGUEZEL; AHMAD; SWAMY, 2009; BOUGUEZEL; AHMAD; SWAMY, 2013; POTLURI et al., 2014; CINTRA; BAYER; TABLADA, 2014). A Tabela 7 apresenta diversos exemplos de matrizes \mathbf{T}_N e \mathbf{S}_N que podem ser encontradas na literatura para o comprimento $N = 8$.

3.2 DCT MULTIDIMENSIONAL APROXIMADA

Embora aproximações para a DCT venham sendo constantemente estudadas e consideradas como ferramentas matemáticas para esquemas de compressão de imagem e vídeo (HAWHEEL, 2001; LENGWEHASATIT; ORTEGA, 2004; CINTRA; BAYER, 2011; BAYER; CINTRA, 2012; BOUGUEZEL; AHMAD; SWAMY, 2008; BOUGUEZEL; AHMAD; SWAMY, 2009; BOUGUEZEL; AHMAD; SWAMY, 2013; POTLURI et al., 2014; CINTRA, 2011; CINTRA; BAYER; TABLADA, 2014), ao nosso melhor conhecimento, não há uma definição formal para o caso geral multidimensional. Nesta seção, objetiva-se propor expressões algébricas baseadas no formalismo tensorial descrito anteriormente. Em seguida, serão propostas aproximações para a 3D DCT baseadas em matrizes aproximadas encontradas na literatura. Também será avaliada a complexidade aritmética para o caso geral multidimensional e para as aproximações 3D obtidas.

3.2.1 Definição Matemática

Visa-se aproximar a DCT multidimensional empregando os vetores de base das aproximações para a DCT encontradas na literatura (HAWHEEL, 2001; LENGWEHASATIT; ORTEGA, 2004; CINTRA; BAYER, 2011; BAYER; CINTRA, 2012; BOUGUEZEL; AHMAD; SWAMY, 2008; BOUGUEZEL; AHMAD; SWAMY, 2009; BOUGUEZEL; AHMAD; SWAMY, 2013; POTLURI et al., 2014). Em geral, tais vetores de base geralmente não apresentam uma expressão matemática explícita, sendo geralmente derivados através de computação numérica e/ou busca exaustiva (CINTRA; BAYER; TABLADA, 2014). Então, em geral, expressões analíticas tais qual (3.1), (3.2) e (3.5) não são convenientes. Entretanto, podemos derivar aproximações multidimensionais para a DCT empregando os conceitos de tensores de alta ordem. Primeiramente consideraremos o caso tridimensional. Seja $\mathcal{X} \in \mathbb{F}^{N \times N \times N}$ um tensor de alta ordem representando um dado sinal discreto de entrada. O tensor de saída no domínio da transformada, $\mathcal{Y} \in \mathbb{F}^{N \times N \times N}$, é dado por

$$\begin{aligned} \mathcal{Y} &\triangleq \mathcal{X} \times_1 \hat{\mathbf{C}}_N \times_2 \hat{\mathbf{C}}_N \times_3 \hat{\mathbf{C}}_N \\ &= \mathcal{X} \times_1 (\mathbf{S}_N \cdot \mathbf{T}_N) \times_2 (\mathbf{S}_N \cdot \mathbf{T}_N) \times_3 (\mathbf{S}_N \cdot \mathbf{T}_N). \end{aligned} \tag{3.22}$$

Empregando a Proposição 3.1 e a Proposição 3.2, podemos reorganizar (3.22) de acordo com o seguinte

$$\mathcal{Y} = \mathcal{X} \times_1 \mathbf{T}_N \times_2 \mathbf{T}_N \times_3 \mathbf{T}_N \times_1 \mathbf{S}_N \times_2 \mathbf{S}_N \times_3 \mathbf{S}_N. \quad (3.23)$$

Então, uma 3D DCT aproximada pode ser obtida computando primeiramente os produtos em modo i -ésimo pelas matrizes de baixa complexidade \mathbf{T}_N . As operações envolvendo a matriz diagonal \mathbf{S}_N podem ser computadas separadamente numa etapa posterior.

A transformação inversa é relacionada à matriz inversa da DCT aproximada, dada em (3.21). Considerando as expressões (3.19) e (3.22), a 3D DCT aproximada inversa é dada por

$$\mathcal{X} = \mathcal{Y} \times_1 \hat{\mathbf{C}}_N^{-1} \times_2 \hat{\mathbf{C}}_N^{-1} \times_3 \hat{\mathbf{C}}_N^{-1}.$$

Pode-se generalizar as expressões anteriores para o caso de RD DCT aproximadas de dimensão $N_1 \times N_2 \times \dots \times N_R$.

Definição 3.2. *Seja $\{\hat{\mathbf{C}}_{N_i}\}_{i=1}^R$ uma coleção de matrizes aproximadas para a DCT de comprimento N_i , em que $\hat{\mathbf{C}}_{N_i} = \mathbf{S}_{N_i} \cdot \mathbf{T}_{N_i}$ para $i = 1, 2, \dots, R$, conforme descrito na Seção 3.1.3. Seja $\mathcal{X} \in \mathbb{F}^{N_1 \times N_2 \times \dots \times N_R}$ um tensor de ordem R -ésima representando um arranjo de entrada. O arranjo de saída no domínio da RD DCT aproximada é definido por*

$$\begin{aligned} \mathcal{Y} &\triangleq \mathcal{X} \times_1 \hat{\mathbf{C}}_{N_1} \times_2 \hat{\mathbf{C}}_{N_2} \times_3 \dots \times_R \hat{\mathbf{C}}_{N_R} \\ &= \mathcal{X} \times_1 (\mathbf{S}_{N_1} \cdot \mathbf{T}_{N_1}) \times_2 (\mathbf{S}_{N_2} \cdot \mathbf{T}_{N_2}) \times_3 \dots \times_R (\mathbf{S}_{N_R} \cdot \mathbf{T}_{N_R}) \\ &= \mathcal{X} \times_1 \mathbf{T}_{N_1} \times_2 \mathbf{T}_{N_2} \times_3 \dots \times_R \mathbf{T}_{N_R} \times_1 \mathbf{S}_{N_1} \times_2 \mathbf{S}_{N_2} \times_3 \dots \times_R \mathbf{S}_{N_R}. \end{aligned} \quad (3.24)$$

Note que a Definição 3.2 implica em uma transformada multidimensional separável, de acordo com Definição 2.2. Note também que, de acordo com tal definição, é possível empregar diferentes aproximações para um comprimento de bloco constante. Dessa forma, se $N_i = N_j$, para $i \neq j$, as aproximações para as i -ésima e j -ésima dimensão não precisam necessariamente ser as mesmas, embora apresentem mesmo comprimento de bloco. Entretanto, selecionar aproximações idênticas para idênticos comprimentos de bloco parece ser uma escolha natural.

Se (i) $N_1 = N_2 = \dots = N_R \triangleq N$, $R \geq 2$, e (ii) a mesma matriz aproximada $\hat{\mathbf{C}}_N$ é considerada em todas as dimensões, a Definição 3.2 implica em uma transformada separável e

Tabela 8 – Eficiência de codificação e complexidade computacional para a 3D DCT exata diversas aproximações para a 3D DCT derivadas

Método	Desempenho	Complexidade 1D			Complexidade 3D		
	C_g (dB)	Mult.	Ad.	Desloc.	Mult.	Ad.	Desloc.
DCT (pela definição)	8.83	64	56	0	12288	10752	0
Algoritmo de Loeffler	8.83	11	29	0	2112	5568	0
Algoritmo de Chen	8.83	16	26	0	3072	4992	0
Algoritmo 3D VR DIF	8.83	–	–	–	1344	5568	0
SDCT	6.03	0	24	0	0	4608	0
LODCT	8.39	0	24	2	0	4608	384
RDCT	8.18	0	22	0	0	4224	0
MRDCT	7.33	0	14	0	0	2688	0
BAS-2008	8.12	0	18	2	0	3456	384
BAS-2009	7.91	0	18	0	0	3456	0
BAS-2013	7.95	0	24	0	0	4608	0
IADCT	7.33	0	14	0	0	2688	0

Fonte: (COUTINHO; CINTRA; BAYER, 2017)

regular, de acordo com a Definição 2.3. A aproximação multidimensional inversa é computada de acordo com a seguinte expressão:

$$\mathcal{X} = \mathcal{Y} \times_1 \hat{\mathbf{C}}_{N_1}^{-1} \times_2 \hat{\mathbf{C}}_{N_2}^{-1} \times_3 \cdots \times_R \hat{\mathbf{C}}_{N_R}^{-1},$$

em que $\hat{\mathbf{C}}_{N_i}^{-1}$ é derivado de (3.21).

3.2.2 Análise de Complexidade

Devido a propriedade de separabilidade de núcleo (KOK, 1997), as DCTs multidimensionais exata e aproximadas podem ser computadas por sucessivas instanciações da DCT ou DCT aproximada (BOUSSAKTA, 2004; SERVAIS; JAGER, 1997; LENGWEHASATIT; ORTEGA, 2004). Conseqüentemente, um algoritmo rápido para a DCT aproximada pode ser aplicado para o cômputo de dimensões maiores. Em (3.24), R diferentes produtos em modo i -ésimo são empregados. Então, aplicando (3.6) em (3.24) para uma dimensão específica i , obtém-se uma expansão com $N_1 \cdot N_2 \cdots N_{i-1} \cdot N_{i+1} \cdots N_R$ índices livres e, então, tal quantidade de produtos em modo i -ésimo. O mesmo raciocínio pode ser empregado às demais dimensões.

Seja $C_{\text{ID}}(\hat{\mathbf{C}}_{N_i})$ a complexidade aritmética de $\hat{\mathbf{C}}_{N_i}$. Tal quantidade engloba os custos multiplicativo, aditivo e de deslocamentos de *bits*, que depende do algoritmo rápido considerado. Sendo assim, a complexidade aritmética para o caso R -dimensional é dado de forma

generalizada por

$$C_{RD} \left(\hat{\mathbf{C}}_{N_1}, \hat{\mathbf{C}}_{N_2}, \dots, \hat{\mathbf{C}}_{N_R} \right) = \sum_{i=1}^R \Omega_R^{(i)} \cdot C_{1D} \left(\hat{\mathbf{C}}_{N_i} \right), \quad (3.25)$$

em que

$$\Omega_R^{(i)} \triangleq \begin{cases} 1, & \text{se } R = 1, \\ \prod_{\substack{j=1 \\ j \neq i}}^R N_j, & \text{caso contrário.} \end{cases}$$

Para o caso 1D, o lado direito de (3.25) se torna simplesmente $C_{1D} \left(\hat{\mathbf{C}}_{N_1} \right)$. Se (i) $N_1 = N_2 = \dots = N_R \triangleq N$, $R \geq 2$, e (ii) a mesma matriz aproximada $\hat{\mathbf{C}}_N$ é considerada em todas as dimensões, então a complexidade aritmética é dada a seguir:

$$C_{RD} \left(\hat{\mathbf{C}}_N \right) = R \cdot N^{R-1} \cdot C_{1D} \left(\hat{\mathbf{C}}_N \right). \quad (3.26)$$

Embora tal abordagem seja generalizada para qualquer comprimento de bloco, focaremos no caso $N = 8$. De fato, tal comprimento de bloco é relevante em diversos contextos práticos (International Telecommunication Union, 2013; LECUIRE; MAKKAOUI; MOUREAUX, 2012; International Organisation for Standardisation, 1994; RAO; HWANG, 1996; WALLACE, 1992).

A Tabela 8 mostra a complexidade computacional para $R = 3$ para as várias aproximações obtidas para a 3D DCT de acordo com o formalismo proposto. Tais aproximações são baseadas nas matrizes aproximadas apresentadas na Tabela 7 e são computadas de acordo com (3.23). Todas as aproximações derivadas apresentam complexidade multiplicativa nula e complexidade aditiva e de deslocamento de *bits* dadas por (3.26). A fim de comparação, também estão inclusas as complexidades computacionais da DCT exata computada (i) pela definição, (ii) pelo algoritmo de Loeffler (LOEFFLER; LIGTENBERG; MOSCHYTZ, 1989), (iii) pelo algoritmo de Chen (CHEN; SMITH; FRALICK, 1977) e (iv) pelo algoritmo 3D VR DIF (BOUSSAKTA, 2004) O algoritmo de Loeffler atinge a cota multiplicativa mínima teórica para o caso 1D (HEIDEMAN, 1988) e o algoritmo de Chen é empregado no padrão HEVC (International Telecommunication Union, 2013). O algoritmo 3D VR DIF (BOUSSAKTA, 2004) computa a 3D DCT exata diretamente sem a necessidade do método de decomposição linha-coluna e apresenta uma menor complexidade multiplicativa do que o algoritmo de Loeffler para o caso 3D

empregando tal decomposição. A Tabela 8 também mostra o ganho de codificação C_g dado em (2.21) como medida de desempenho.

A Tabela 9 mostra as reduções percentuais nas métricas de performance de codificação e complexidade computacional para cada aproximação para a 3D DCT derivada comparada com o método exato empregando o algoritmo 3D VR DIF. A redução total de complexidade aritmética é obtida considerando a soma de todas as operações aritméticas mostradas na Tabela 8. As reduções percentuais de complexidade apresentadas pelas aproximações propostas são mais acentuadas do que as reduções percentuais em desempenho. Este fato sugere um *trade-off* favorável entre complexidade e desempenho. Os métodos 3D MRDCT e 3D IADCT igualmente apresentam redução de 61,1% e 51,7% no custo aritmético total e custo aditivo, respectivamente, quando comparados com o algoritmo 3D VR DIF. O método 3D LODCT mostra a menor degradação de desempenho, com reduções de 4,9% no ganho de codificação, respectivamente, a uma economia de 27,8% no custo computacional total.

3.2.3 Análise de *Trade-off*

O desempenho geral de uma particular aproximação para a 3D DCT em um contexto específico depende de diversos fatores de *trade-off* (BRIGGS; HENSON, 1995), tais como consumo de potência, área de *chip* ocupada, uso de memória, estabilidade numérica, nível de paralelismo, latência e máxima frequência de operação, não sendo uma tarefa trivial os quantificar precisamente (LECUIRE; MAKKAOUI; MOUREAUX, 2012; CINTRA et al., 2016). No entanto, uma análise simplificada pode ser realizada através de um problema multiobjetivo empregando uma figura de mérito f que leve em consideração complexidade computacional e desempenho de codificação, que são duas importantes métricas na área (BLAHUT, 2010; BRITANAK; YIP; RAO, 2007). Sendo assim, propõe-se a seguinte combinação convexa como figura de mérito para analisar as aproximações propostas para a 3D DCT:

$$f \triangleq \gamma \cdot (\text{complexidade computacional normalizada}) + (1 - \gamma) \cdot (\text{desempenho normalizado}), \quad (3.27)$$

em que $\gamma \in [0, 1]$ é fator de ponderação e a normalização das medidas é considerada para garantir que os valores sejam comparáveis e é tomada relativa ao maior valor entre todos os

métodos. A quantidade γ ajusta a importância relativa de cada componente — complexidade computacional e desempenho — de acordo com um determinado contexto.

Para a métrica de complexidade computacional em (3.27), é adotado o custo aritmético, que definimos pela soma ponderada das complexidades multiplicativa, aditiva e de deslocamentos de *bits*. Seja M , A , e S o número de operações de multiplicações, adições e deslocamentos de *bits*, respectivamente, e w_m , w_a e w_s seus respectivos pesos para o custo computacional total. Então, definimos

$$\begin{aligned} (\text{custo aritmético}) &\triangleq w_m \cdot M + w_a \cdot A + w_s \cdot S \\ &= w_m \cdot \left(M + \frac{w_a}{w_m} \cdot A + \frac{w_s}{w_m} \cdot S \right) \\ &= w_m \cdot (M + \beta \cdot A + \beta' \cdot S), \end{aligned} \tag{3.28}$$

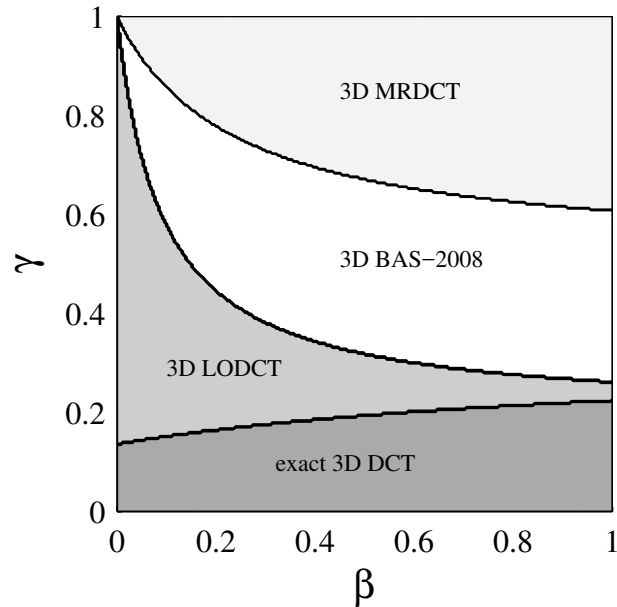
em que $\beta = w_a/w_m$ e $\beta' = w_s/w_m$. Entretanto, dado que a operação de deslocamento de *bits* requer consideravelmente menos consumo de energia e recursos de *hardware* quando comparado com operações de adição e multiplicação, consideramos que $w_m \gg w_s$; então, obtém-se $\beta' \approx 0$. Ainda mais, dado que (3.27) considera a complexidade computacional normalizada, o termo w_m não tem efeito prático. Doravante, por simplicidade, consideramos $w_m = 1$. Consequentemente, de (3.28) obtém-se

$$(\text{custo arimético}) \approx M + \beta A. \tag{3.29}$$

Atribuir um valor fixo para β não é trivial (BRIGGS; HENSON, 1995, p. 395). Entretanto, é um fato bem difundido que operações de multiplicação são mais complexas do que adições ($w_m > w_a$) em implementações tanto em *hardware* quanto em *software* (BLAHUT, 2010). Então, tem-se que $\beta < 1$. Além disso, sistemas computacionais tendem a satisfazer $w_a \ll w_m$; logo valores baixos para β são esperados em sistemas práticos. Sendo assim, adotamos $\beta \in [0, 1]$ em nossa análise. Por sua vez, para a métrica de desempenho em (3.27), adotamos o negativo do ganho de codificação, tal que a figura de mérito decresce quando a performance de codificação aumenta (relação inversa entre custo e desempenho).

Todas as aproximações para a 3D DCT derivadas assim como a 3D DCT exata via algoritmo 3D VR DIF foram submetidas à figura de mérito proposta para todos os valores de γ e β . Na Figura 12, destacam-se os métodos 3D ótimos de acordo com a função custo proposta para

Figura 12 – Transformadas 3D ótimas para diferentes parâmetros da função custo.



Fonte: (COUTINHO; CINTRA; BAYER, 2017)

cada região no plano $\beta\gamma$. Valores baixos de γ enfatizam performance de codificação na função custo. Então, conforme esperado, a 3D DCT exata tende a ser o método ótimo. A partir de $\gamma \approx 0.2$, as aproximações superam a 3D DCT exata. Para valores de γ médio-baixo, médio-alto e alto, a 3D LODCT, a 3D BAS-2008 e a 3D MRDCT são os métodos ótimos, respectivamente. Agora, analisemos o eixo β . Tem-se para valores baixos de β — que enfatizam complexidade multiplicativa sobre a aditiva — que a 3D LODCT tende a superar as demais aproximações. Incrementando β , o que dá mais equilíbrio entre os custos relativos entre operações de multiplicação e adição para o custo aritmético total, tem-se a 3D BAS-2008 e a 3D MRDCT ocupando uma maior área da região. Além disso, a área em que a 3D DCT exata é o método ótimo cresce ligeiramente na medida em que β incrementa. Em geral, as aproximações para a 3D DCT derivadas ocupam uma área maior no plano, superando a 3D DCT em situações realísticas.

3.3 CODIFICAÇÃO DE VÍDEO ENTRE QUADROS BASEADA EM APROXIMAÇÕES PARA A 3D DCT

A presente seção propõe um esquema de codificação de vídeo baseado em transformada de bloco 3D (SERVAIS; JAGER, 1997; NATARAJAN; AHMED, 1977; LAI; GUAN, 2002; LI;

Tabela 9 – Reduções percentuais de eficiência de codificação e complexidade computacional comparado com a 3D DCT exata (levando em conta o algoritmo 3D VR DIF)

Método	Desempenho		Complexidade 3D	
	C_g	Multiplicações	Adições	Total
3D SDCT	31.7 %	100 %	17.2 %	33.3 %
3D LODCT	4.9 %	100 %	17.2 %	27.8 %
3D RDCT	7.3 %	100 %	24.1 %	38.9 %
3D MRDCT	16.9 %	100 %	51.7 %	61.1 %
3D BAS-2008	8.0 %	100 %	37.9 %	44.4 %
3D BAS-2009	10.4 %	100 %	37.9 %	50.0 %
3D BAS-2013	10.0 %	100 %	17.2 %	33.3 %
3D IADCT	16.9 %	100 %	51.7 %	61.1 %

Fonte: (COUTINHO; CINTRA; BAYER, 2017)

HOU, 2007) que emprega as aproximações para a 3D DCT propostas (veja Tabela 7 e (3.23), pp. 66 e 67). As aproximações para a 3D DCT são embutidas em um algoritmo de codificação de vídeo que se baseia na 3D DCT exata (NATARAJAN; AHMED, 1977; SERVAIS; JAGER, 1997; ABOUSLEMAN; MARCELLIN; HUNT, 1995; LAI; GUAN, 2002; LI; HOU, 2007). Também é proposto um método para modificar um dado volume de quantização 3D de modo a evitar multiplicações extras nas aproximações propostas. Então, um conjunto de sequências de vídeo padrão é submetido ao sistema de codificação de vídeo modificado para avaliar o desempenho em termos de qualidade de vídeo.

3.3.1 Etapa de Quantização Modificada para Compressão de Vídeo

Em esquemas de compressão de imagem e vídeo, a etapa de quantização exerce um papel importante uma vez que descarta não linearmente componentes no domínio da frequência de acordo com (i) suas respectivas importâncias para a qualidade visual percebida (BHASKARAN; KONSTANTINIDES, 1997, p. 155) ou (ii) propriedades de correlação dos dados (BHASKARAN; KONSTANTINIDES, 1997, p. 156). Em esquemas de codificação usuais baseados na 2D DCT, o procedimento de quantização depende de tabelas (ou matrizes) de quantização (RAO; YIP, 2001; BHASKARAN; KONSTANTINIDES, 1997), que são adotadas pelos principais codecs, tais como MPEG (WALLACE, 1992; International Organisation for Standardisation, 1994), H.261 (International Telecommunication Union, 1990) H.263 (International Telecommunication Union, 1995), H.264 (TEAM, 2003) e HEVC (International Telecommu-

nication Union, 2013).

Como detalhado em (BRITANAK; YIP; RAO, 2007; TRAN, 2000; LIANG; TRAN, 2001; LENGWEHASATIT; ORTEGA, 2004), a matriz diagonal da DCT aproximada (3.20) pode ser incorporada à tabela de quantização, eliminando a necessidade de operações adicionais implicadas por tal matriz diagonal. Entretanto, para esquemas de codificação baseados na 3D DCT, tais tabelas de quantização não se adequam por serem inerentemente arranjos bidimensionais (LEE; CHAN; ADJEROH, 1997; CHAN; LEE, 1997; BOZINOVIĆ; KONRAD, 2003; BOZINOVIĆ; KONRAD, 2005). No lugar de uma tabela de quantização, um volume de quantização se faz necessário. Em (LEE; CHAN; ADJEROH, 1997; CHAN; LEE, 1997; BOZINOVIĆ; KONRAD, 2003), são propostos métodos para gerar volumes de quantização. Na presente seção, visa-se propor um método para embutir as matrizes diagonais descritas em (3.23) em um dado volume de quantização arbitrário.

Seja $\mathcal{Q} \in \mathbb{F}^{N \times N \times N}$ um volume de quantização previamente construído, cujas entradas são dadas por $q[k_1, k_2, k_3]$, para $k_1, k_2, k_3 = 0, 1, \dots, N - 1$. O bloco de quantização realiza o seguinte cômputo (LEE; CHAN; ADJEROH, 1997; CHAN; LEE, 1997):

$$\tilde{y}[k_1, k_2, k_3] = \text{round} \left(\frac{y[k_1, k_2, k_3]}{q[k_1, k_2, k_3]} \right), \quad (3.30)$$

$$k_1, k_2, k_3 = 0, 1, \dots, N - 1,$$

em que $y[k_1, k_2, k_3]$ são os coeficientes no domínio da transformada de acordo com (3.23), $\text{round}(\cdot)$ é a função de arredondamento para o inteiro mais próximo e $\tilde{y}[k_1, k_2, k_3]$ são os coeficientes quantizados. O processo de desquantização é definido por $\hat{y}[k_1, k_2, k_3] = \tilde{y}[k_1, k_2, k_3] \cdot q[k_1, k_2, k_3]$, em que $\hat{y}[k_1, k_2, k_3]$ é uma estimativa de $y[k_1, k_2, k_3]$ (BHASKARAN; KONSTANTINIDES, 1997).

Proposição 3.3. *Seja o cálculo da 3D DCT aproximada dada em (3.23) separado em dois passos: (i) o cômputo dos produtos em modo i -ésimo envolvendo a matriz de baixa complexidade \mathbf{T} , dado por*

$$\mathcal{A} = \mathcal{X} \times_1 \mathbf{T}_N \times_2 \mathbf{T}_N \times_3 \mathbf{T}_N, \quad (3.31)$$

e (ii) os produtos em modo i -ésimo envolvendo a matriz diagonal \mathbf{S}_N . Então, é válida a seguinte

relação:

$$y[k_1, k_2, k_3] = a[k_1, k_2, k_3] \cdot d_{k_1} \cdot d_{k_2} \cdot d_{k_3}, \quad (3.32)$$

em que $a[k_1, k_2, k_3]$ são as entradas do tensor \mathcal{A} dado em (3.31) e d_k é o k -ésimo elemento diagonal de \mathbf{S}_N .

Demonstração. Para obter a 3D DCT aproximada dada em (3.23) a partir do tensor \mathcal{A} dado por (3.31), seja o tensor intermediário

$$\mathcal{B} = \mathcal{A} \times_1 \mathbf{S}_N.$$

Sejam as entradas da matriz diagonal \mathbf{S}_N

$$s[k, n] \triangleq \begin{cases} d_k, & \text{se } k = n, \\ 0, & \text{caso contrário,} \end{cases} \quad n, k = 0, 1, \dots, N-1. \quad (3.33)$$

Considerando (3.6), as entradas do tensor \mathcal{B} são dadas por

$$b[k_1, n_2, n_3] = \sum_{n_1=0}^{N-1} a[n_1, n_2, n_3] \cdot s[k_1, n_1]. \quad (3.34)$$

Substituindo (3.33) em (3.34), obtém-se

$$b[k_1, n_2, n_3] = a[k_1, n_2, n_3] \cdot d_{k_1}.$$

Agora, seja o próximo tensor intermediário dado por

$$\mathcal{C} = \mathcal{B} \times_2 \mathbf{S}_N.$$

De maneira similar, as entradas do tensor \mathcal{C} são dadas a seguir:

$$\begin{aligned} c[k_1, k_2, n_3] &= b[k_1, k_2, n_3] \cdot d_{k_2} \\ &= a[k_1, k_2, n_3] \cdot d_{k_1} \cdot d_{k_2}. \end{aligned}$$

Por fim, o tensor de saída \mathcal{Y} dado em (3.23) pode ser expresso por

$$\mathcal{Y} = \mathcal{C} \times_3 \mathbf{S}_N,$$

e suas entradas são dadas, analogamente, por

$$\begin{aligned} y[k_1, k_2, k_3] &= c[k_1, k_2, k_3] \cdot d_{k_3} \\ &= a[k_1, k_2, k_3] \cdot d_{k_1} \cdot d_{k_2} \cdot d_{k_3}. \end{aligned}$$

□

Empregando a Proposição 3.3 e aplicando (3.32) em (3.30), obtém-se

$$\tilde{y}[k_1, k_2, k_3] = \text{round} \left(\frac{a[k_1, k_2, k_3] \cdot d_{k_1} \cdot d_{k_2} \cdot d_{k_3}}{q[k_1, k_2, k_3]} \right), \quad (3.35)$$

$$k_1, k_2, k_3 = 0, 1, \dots, N - 1.$$

Definição 3.3. *Seja $\mathcal{Q} \in \mathbb{F}^{N \times N \times N}$ um volume de quantização previamente construído, defina-se o novo volume de quantização modificado $\tilde{\mathcal{Q}} \in \mathbb{F}^{N \times N \times N}$, cujas entradas são dadas por*

$$\tilde{q}[k_1, k_2, k_3] \triangleq \frac{q[k_1, k_2, k_3]}{d_{k_1} \cdot d_{k_2} \cdot d_{k_3}}. \quad (3.36)$$

Empregando a Definição 3.3 e substituindo (3.36) em (3.35), obtém-se a etapa de quantização modificada proposta, conforme mostrado a seguir:

$$\tilde{y}[k_1, k_2, k_3] = \text{round} \left(\frac{a[k_1, k_2, k_3]}{\tilde{q}[k_1, k_2, k_3]} \right), \quad (3.37)$$

$$k_1, k_2, k_3 = 0, 1, \dots, N - 1.$$

Note que empregando o volume de quantização modificado proposto, é necessário computar apenas o tensor \mathcal{A} que é aplicado como tensor de entrada ao bloco de quantização modificado (3.37). Dado que o cômputo de \mathcal{A} requer apenas as transformações de baixa complexidade \mathbf{T}_N de acordo com (3.31), a sobrecarga computacional imposta pela matriz diagonal é eliminada. O procedimento de desquantização modificado pode ser obtido de forma análoga e é dado por

$$\hat{y}[k_1, k_2, k_3] = \tilde{y}[k_1, k_2, k_3] \cdot \tilde{q}_{\text{inv}}[k_1, k_2, k_3],$$

em que

$$\tilde{q}_{\text{inv}}[k_1, k_2, k_3] \triangleq q[k_1, k_2, k_3] \cdot d_{k_1} \cdot d_{k_2} \cdot d_{k_3}.$$

3.3.2 Simulação de Compressão de Vídeo

Nesta seção, um esquema de compressão de vídeo baseado em (NATARAJAN; AHMED, 1977; SERVAIS; JAGER, 1997; ABOUSLEMAN; MARCELLIN; HUNT, 1995; LAI; GUAN, 2002; LI; HOU, 2007) foi considerado. Um conjunto padrão de nove sequências de vídeos no formato CIF YUV (FITZEK; REISSLEIN, 2001) disponível em (XIPH. . .) foi considerado, em que os vídeos modelados como tensores de terceira ordem de dimensão $N_1 \times N_2 \times$

N_3 . Para o formato CIF, $N_1 = 352$ e $N_2 = 288$; e foram selecionados 296 quadros de vídeo consecutivos ($N_3 = 296$).

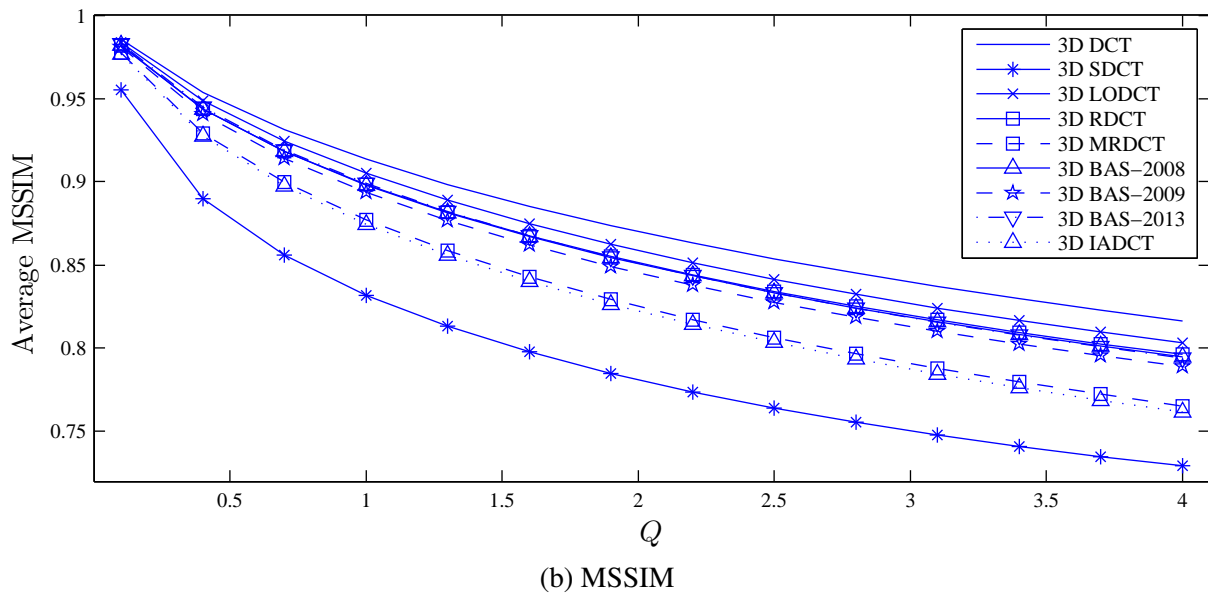
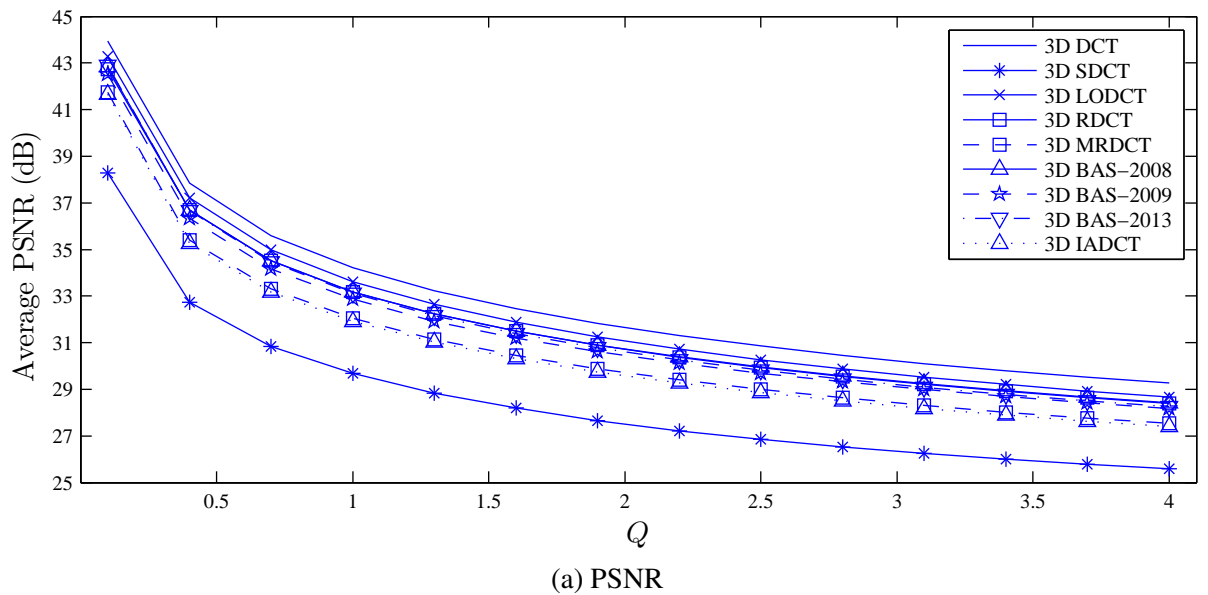
Cada tensor de vídeo foi dividido em blocos tensores menores de dimensões $8 \times 8 \times 8$, que foram então aplicados a cada transformação 3D definida em (3.23) e depois ao estágio de quantização. As aproximações para a 3D DCT mostradas na Tabela 7 foram consideradas. Adicionalmente, a 3D DCT exata também foi incluída nos experimentos para comparação. O volume de quantização sugerido em (LEE; CHAN; ADJEROH, 1997) foi empregado para gerar o bloco de quantização modificado sugerido na Seção 3.3.1. O algoritmo de compressão de vídeo foi simulado variando o fator de qualidade $Q \in [0.1, 4.0]$ (BHASKARAN; KONSTANTINIDES, 1997). O procedimento inverso é aplicado para reconstruir a sequência de vídeo. Tal procedimento inverso engloba a desquantização modificada e a transformação 3D inversa.

Para avaliar o desempenho de compressão, emprega-se as métricas de PSNR e SSIM descritas na Seção 2.3.1. Uma vez que o SSIM é definido para imagens estáticas (WANG et al., 2004), considera-se o SSIM médio obtido entre quadros de vídeo. Os valores médios de PSNR e MSSIM em função de Q entre as sequências de vídeo consideradas são apresentados na Figura 13. A 3D DCT exata apresenta a maior qualidade de vídeo, enquanto a 3D SDCT mostra o pior desempenho comparado com as demais aproximações para a 3D DCT. Entretanto, a maioria dos métodos mostram desempenho similar a 3D DCT exata, especialmente a baixas taxas de compressão (valores pequenos de Q).

Na análise a seguir, excluímos a 3D DCT uma vez que os demais métodos aproximados concorrentes apresentam um desempenho consideravelmente mais alto. De forma geral, os métodos aproximados apresentam menos de 7% de degradação em PSNR comparados a 3D DCT exata. Por exemplo, para $Q = 0.1$, a 3D LODCT apresenta PSNR apenas 1.53% menor quando comparado com a 3D DCT exata, enquanto a 3D IADCT mostra uma redução de PSNR de 5.22%. Para $Q = 4$, tais valores são 2.01% e 6.36%, respectivamente. A Figura 14 mostra uma comparação qualitativa para o primeiro quadro da sequência de vídeo “foreman” para $Q = 0.25$. Os quadros são quase indistinguíveis.

Também foi averiguado o impacto do nível de movimento (*motion level*) de cada sequência no desempenho de codificação para cada um dos métodos propostos. De acordo com (DAWOOD;

Figura 13 – Medidas de desempenho em compressão de vídeo para a 3D DCT exata e diversas aproximações 3D derivadas.



Fonte: (COUTINHO; CINTRA; BAYER, 2017)

GHANBARI, 1999), o nível de movimento pode ser quantificado empregando vetores de movimento m , obtidos empregando algoritmos de estimação de movimento. Como métrica de nível de movimento, adota-se a magnitude média dos vetores de movimento $|\overline{m}|$ (DAWOOD; GHANBARI, 1999). Os vetores de movimento foram extraídos dos quadros de vídeo empregando o algoritmo ARPS de estimação de movimento (NIE; MA, 2002; BARJATYA, 2004). Na Tabela 10, mostra-se o nível de movimento de cada sequência de vídeo considerada em ordem

Figura 14 – Comparação qualitativa para o primeiro quadro da sequência de vídeo “foreman” empregando a 3D DCT exata e diversas aproximações 3D propostas.



(a) Original



(b) 3D DCT (PSNR=37.8988, MSSIM = 0.9456), (c) 3D SDCT (PSNR=32.3627, MSSIM = 0.8519), (d) 3D LODCT (PSNR=37.2166, MS-SIM = 0.9393)



(e) 3D RDCT (PSNR=36.7865, MSSIM = 0.9352), (f) 3D MRDCT (PSNR=36.5764, MSSIM = 0.9280), (g) BAS-2008 (PSNR=37.0146, MSSIM = 0.9378)



(h) 3D BAS-2009 (PSNR=36.5367, MS-SIM = 0.9341), (i) 3D BAS-2013 (PSNR=36.3304, MSSIM = 0.9323), (j) 3D IADCT (PSNR=35.2330, MSSIM = 0.9135)

crecente.

É esperado que a 3D DCT apresente um menor desempenho para vídeos com maior nível de movimento, uma vez que uma alta variação temporal de *pixels* sugere uma dispersão de energia maior nos coeficientes da 3D DCT. Entretanto, outros fatores também influenciam na qualidade do vídeo comprimido, tal qual o nível de textura espacial dos quadros (DAWOOD; GHANBARI, 1999).¹ Uma vez que as aproximações para a 3D DCT apresentam propriedades similares ao caso exato por construção, espera-se que as mesmas demonstrem o mesmo comportamento em relação ao nível de movimento das sequências de vídeo.

Na Tabela 11, é mostrado o desempenho de cada transformação 3D para cada sequência de vídeo, assim como o coeficiente de correlação de Pearson ρ (PEARSON, 1901; BENESTY et al., 2009) entre a métrica de desempenho e o nível de movimento. De fato, os valores de correlação mostram uma tendência inversa entre desempenho e nível de movimento, especialmente para a métrica de SSIM que melhor captura a qualidade visual percebida pelo olho humano (WANG et al., 2004). Além disso, todos os métodos aproximados apresentam valores de correlação semelhantes ao da 3D DCT exata, sugerindo que as propriedades de compactação são preservadas. Sendo assim, os métodos 3D propostos são apropriados para substituir a 3D DCT exata em sistemas de codificação de vídeo independentemente do nível de movimento considerado.

Tabela 10 – Nível de movimento das sequências de vídeos consideradas

Índice da sequência	Sequência	Nível de movimento $ \overline{\mathbf{m}} $
1	akyio	0.0387
2	container	0.0954
3	news	0.2561
4	silent	0.4877
5	mobile	0.748
6	mother-daughter	0.7594
7	Hall-monitor	0.805
8	coastguard	1.7614
9	foreman	2.345

Fonte: (COUTINHO; CINTRA; BAYER, 2017)

¹ Embora textura seja característica básica de uma imagem e que pode ser entendida de forma intuitiva, não há uma definição formal para tal qualidade (BHARATI; LIU; MACGREGOR, 2004). Diferentes métodos vêm sendo propostos para analisar e quantificar a textura de uma imagem (DONG et al., 2017).

Tabela 11 – Desempenho dos métodos propostos para a 3D DCT para cada sequência de vídeo e correlação com nível de movimento

Medida	Método	Índice da sequência									correlação ρ
		1	2	3	4	5	6	7	8	9	
PSNR (dB)	3D DCT	39.12	35.87	36.33	34.93	26.42	37.74	35.68	30.22	31.69	-0.55
	3D SDCT	34.31	31.41	31.31	30.74	21.05	33.27	32.16	25.77	27.08	-0.51
	3D LODCT	38.27	35.21	35.59	34.42	25.86	37.07	35.24	29.79	31.18	-0.54
	3D RDCT	37.80	34.83	35.06	34.06	25.11	36.66	34.95	29.36	30.82	-0.52
	3D MRDCT	36.71	33.68	33.88	33.13	23.42	35.58	34.28	28.13	29.61	-0.51
	3D BAS-2008	37.88	34.76	35.22	34.12	24.94	36.68	35.06	29.24	30.75	-0.53
	3D BAS-2009	37.52	34.50	34.83	33.87	24.50	36.34	34.81	28.96	30.44	-0.52
	3D BAS-2013	37.69	34.73	35.03	34.10	25.14	36.50	34.90	29.29	30.63	-0.54
	3D IADCT	36.56	33.53	33.76	32.99	23.26	35.41	34.17	28.02	29.51	-0.50
	SSIM	3D DCT	0.96	0.93	0.95	0.92	0.87	0.94	0.93	0.86	0.86
3D SDCT		0.92	0.89	0.90	0.84	0.72	0.87	0.89	0.74	0.73	-0.78
3D LODCT		0.95	0.93	0.94	0.91	0.86	0.93	0.93	0.85	0.84	-0.84
3D RDCT		0.95	0.92	0.94	0.91	0.84	0.92	0.92	0.84	0.84	-0.82
3D MRDCT		0.94	0.91	0.93	0.89	0.79	0.91	0.92	0.80	0.80	-0.76
3D BAS-2008		0.95	0.92	0.94	0.91	0.84	0.93	0.93	0.83	0.84	-0.80
3D BAS-2009		0.95	0.92	0.94	0.91	0.83	0.92	0.92	0.82	0.83	-0.80
3D BAS-2013		0.95	0.92	0.94	0.91	0.84	0.92	0.92	0.83	0.83	-0.84
3D IADCT		0.94	0.91	0.93	0.89	0.78	0.91	0.92	0.79	0.80	-0.76

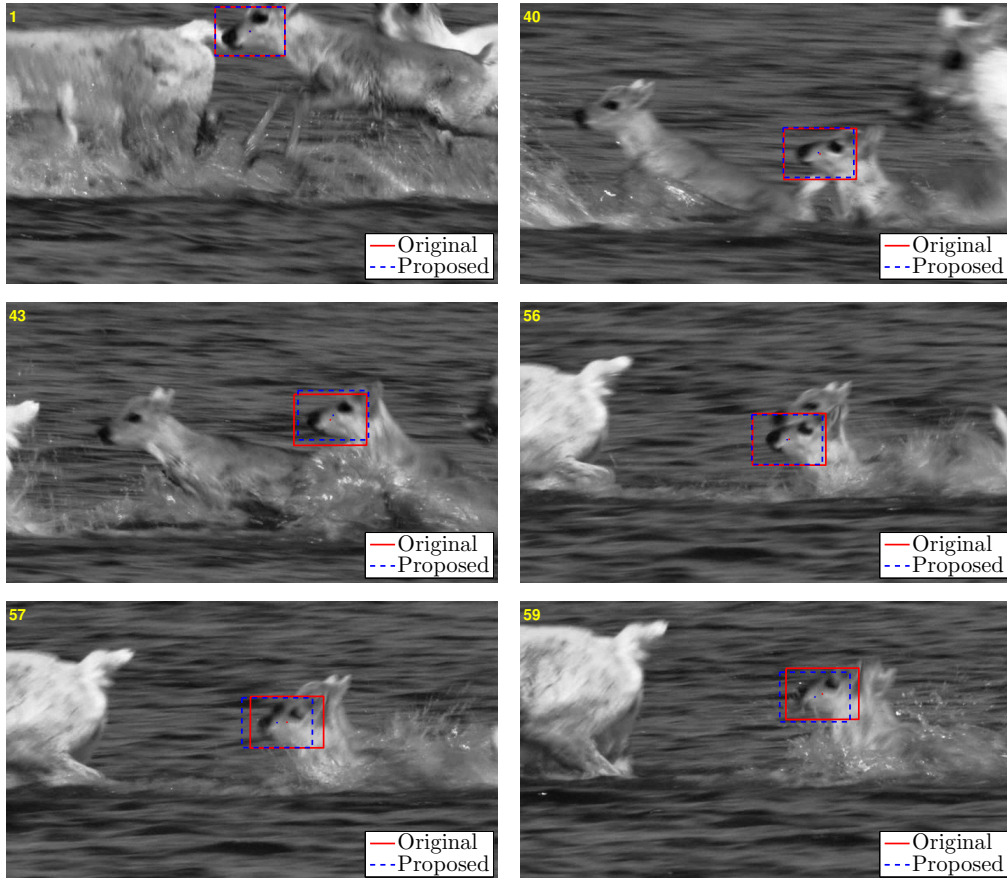
Fonte: (COUTINHO; CINTRA; BAYER, 2017)

3.4 SISTEMA DE RASTREAMENTO VISUAL BASEADO EM UMA APROXIMAÇÃO PARA A 3D DCT

Rastreamento visual consiste em prever a localização de um alvo a partir de uma posição inicial em uma sequência de quadros de vídeo. Uma grande variedade de métodos pode ser encontrada na literatura (SMEULDERS et al., 2014). Aplicações diretas envolvem: vigilância visual e controle de segurança (COIFMAN et al., 1998; JAVED; SHAH, 2002; WANG; WANG; BI, 2009), assistência de motorista e rastreamento de veículos (COIFMAN et al., 1998; ALMAGAMBETOV; VELIPASALAR; CASARES, 2015) e redes de sensores visuais sem fio (WSVN) (WANG; WANG; BI, 2009; DEMIGHA; HIDOUICI; AHMED, 2013). A classe de rastreadores discutida em (SMEULDERS et al., 2014; ROSS et al., 2008; ABDI; WILLIAMS, 2010) emprega análise de componentes principais (PCA) como modelo de aparência (SMEULDERS et al., 2014; ROSS et al., 2008). Matematicamente, A PCA é equivalente a KLT (GERBRANDS, 1981).

Tomando como base rastreadores baseados em PCA e a relação assintótica entre a KLT e a DCT, foi proposto um sistema de rastreamento visual baseado na 3D DCT para representação

Figura 15 – Avaliação qualitativa para rastreamento da sequência de vídeo “animal” empregando a 3D DCT exata e a aproximação 3D MRDCT para alguns quadros representativos (quadros de número 1, 40, 43, 56, 57 e 59).

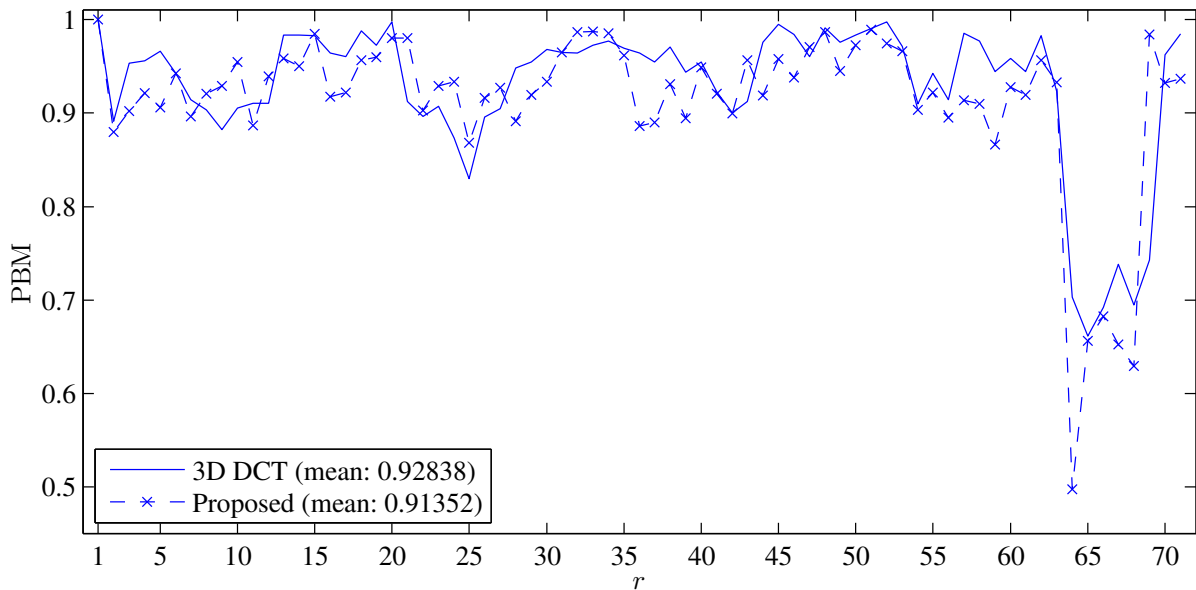


Fonte: (COUTINHO; CINTRA; BAYER, 2017)

de subespaço em menor dimensão. Pelas vantagens computacionais da DCT em comparação com a KLT, tal algoritmo sugere uma implementação mais eficiente em termos de custo computacional quando comparada com os rastreadores baseados em PCA (LI et al., 2013; ROSS et al., 2008).

Algoritmos de baixa complexidade para sistemas de rastreamento visual beneficiam aplicações que requerem processamento em tempo real (CHANG; LIE, 2012; ALMAGAMBE-TOV; VELIPASALAR; CASARES, 2015; ZHOU et al., 2015). Como discutido em (BRITANAK; YIP; RAO, 2007; LIANG; TRAN, 2001; LENGWEHASATIT; ORTEGA, 2004; CINTRA; BAYER, 2011), aproximações para a DCT são ferramentas de baixo custo computacional que vêm sendo desenvolvidas para tecnologias que se baseiam na DCT. Na presente seção, emprega-se uma aproximação para a 3D DCT em um sistema de rastreamento visual.

Figura 16 – Métrica PBM para cada r -ésimo quadro da sequência de vídeo “animal” empregando a 3D DCT exata e a aproximação 3D MRDCT.



Fonte: (COUTINHO; CINTRA; BAYER, 2017)

O bloco da 3D DCT no sistema proposto (LI et al., 2013) é modificado de modo a computar uma aproximação para a 3D DCT. O sistema original computa a 3D DCT exata de dimensões $N_1 \times N_2 \times N_3$ através de um algoritmo incremental que explora uma transformada rápida de Fourier (AHMED; NATARAJAN; RAO, 1974). Os valores N_1 e N_2 foram mantidos fixos, enquanto N_3 varia incrementalmente até um tamanho T máximo de *buffer*. O método demanda $N_1 N_2 (\log_2 N_1 + \log_2 N_2) + N_1 N_2 N_3 \log_2 N_3$ multiplicações complexas e o dobro de tal valor de adições complexas (LI et al., 2013; BLAHUT, 2010). Fixa-se $N_1 = N_2 = T = 8$ de maneira a compatibilizar com o comprimento de bloco das aproximações derivadas no presente trabalho na Seção 3.2.

Dentre as aproximações apresentadas na Tabela 7, selecionamos a 3D MRDCT, que apresenta menor custo aditivo conforme mostrado na Tabela 8. A 3D IADCT apresenta um mesmo custo computacional, porém um menor desempenho no algoritmo de compressão de vídeo. Para os quadros de vídeo iniciais, em que $N_3 < T$, emprega-se uma abordagem híbrida aproximada/exata através de (3.24) com $R = 3$. Utiliza-se a matriz da MRDCT para \hat{C}_{N_1} e \hat{C}_{N_2} e a matriz da DCT exata C_{N_3} no lugar de \hat{C}_{N_3} até que N_3 atinja o valor final $N_3 = T = 8$. Então, computa-se a 3D MRDCT completa conforme proposto em (3.22) e (3.23) para os

quadros restantes.

Neste caso, o sistema rastreador original exige 1920 e 3840 multiplicações e adições complexas, respectivamente, enquanto o rastreador modificado requer apenas 2688 adições reais, conforme mostrado na Tabela 8. Todos os demais parâmetros do sistema de rastreamento original e modificado foram mantidos constantes.

É válido salientar que o principal propósito desta seção é prover uma prova-de-conceito para a 3D DCT aproximada; sugerindo que aproximações de baixa complexidade são alternativas viáveis para o cômputo de tais tipos de sistemas de rastreamento visual. Sendo assim, consideramos o sistema original proposto em (LI et al., 2013) como referência para análise. A Figura 15 mostra uma comparação qualitativa para alguns quadros representativos da sequência de vídeo ‘animal’ disponível em (COMPUTER...). Ambos os sistemas original e modificado apresentam desempenho próximo.

Para uma avaliação quantitativa foi computada a métrica baseada em posição — ou *position-based measure* (PBM) (SMEULDERS et al., 2014). Tal medida é baseada na distância do centróide da região-alvo relativa ao centróide da região verdadeira previamente definida (*ground truth*). Empregamos o *ground truth* provido em (COMPUTER...). Seja \mathbb{T}_r e \mathbb{G}_r as regiões retangulares (*bounding boxes*) do alvo e do *ground truth*, respectivamente, para o r -ésimo quadro de vídeo. A medida PBM é dada por (SMEULDERS et al., 2014)

$$\text{PBM}(r) = 1 - \frac{D(r)}{T_h(r)}, \quad (3.38)$$

em que $T_h(r) = [\text{width}(\mathbb{T}_r) + \text{height}(\mathbb{T}_r) + \text{width}(\mathbb{G}_r) + \text{height}(\mathbb{G}_r)]/2$ e

$$D(r) = \begin{cases} \|\text{cent}(\mathbb{T}_r) - \text{cent}(\mathbb{G}_r)\|, & \text{se } \mathbb{G}_r \cap \mathbb{T}_r \neq \emptyset, \\ T_h(r), & \text{caso contrário;} \end{cases}$$

e $\text{cent}(\cdot)$, $\text{width}(\cdot)$ e $\text{height}(\cdot)$ retornam o centróide, o comprimento e a altura de seus argumentos, respectivamente. Os valores de PBM são confinados no intervalo $[0, 1]$. O valor 0 representa uma falha de rastreamento, enquanto valor 1 indica que os centróides de \mathbb{T}_r e \mathbb{G}_r são os mesmos. Sendo assim, valores próximos a 1 sugerem bom desempenho de rastreamento. A Figura 16 mostra o PBM para todos os quadros r da sequência de vídeo “animal”. Os valores para o método modificado proposto tendem a seguir a curva do método original, sendo algumas

vezes até maior. Em média, o método proposto apresenta um valor de PBM apenas 1.6% menor e a um custo computacional de transformada 3D consideravelmente menor, com reduções de 100 % e 52 % de complexidade multiplicativa e aditiva, respectivamente.

4 APROXIMAÇÕES PARA A DHT TRIDIMENSIONAL E APLICAÇÕES EM COMPRESSÃO DE IMAGEM MÉDICA

A transformada integral proposta por Hartley em (HARTLEY, 1942) foi introduzida como uma alternativa para análise de sinais via transformada de Fourier. Tal transformada apresenta a vantagem de mapear sinais reais em um domínio espectral também real, diferentemente da transformada de Fourier clássica. Além disso, a transformada introduzida representa uma involução, isto é, as transformadas direta e inversa são idênticas, a menos de um fator de escala.

Com a crescente demanda por sistemas digitais, que impulsionou o desenvolvimento de transformadas discretas (AHMED; RAO, 1975), a DHT surge como uma versão discreta para a transformada integral proposta por Hartley (BRACEWELL, 1983) e que, de forma similar ao caso contínuo, apresenta uma relação direta com a DFT. De fato, foi determinado em (HEIDEMAN, 1988, p. 115) que a DHT e a DFT são sistemas equivalentes em termos de complexidade multiplicativa. Conseqüentemente, ambas as transformadas apresentam a mesma cota inferior para a complexidade multiplicativa quando considerado o mesmo comprimento de transformada. No entanto, a DHT apresenta propriedades de involução e espectro real análogas ao caso contínuo (BRACEWELL, 1983; GONZALEZ; WOODS, 2001), que representam vantagens da DHT em relação a DFT em termos de implementação (BRACEWELL, 1984) — particularmente em aplicações óticas (BRACEWELL et al., 1985; CHEN; ZHAO, 2006; MOREOLO; MUÑOZ; JUNYENT, 2010; SINGH; SINHA, 2009; ZHAO; LI; CHEN, 2008). Ainda mais, a DHT é também uma alternativa para computar outras transformadas sinusoidais, tais como a DCT (MALVAR, 1986) e a transformada discreta do seno (DST) (MEHER; SRIKANTHAN; PATRA, 2006).

Além de ser considerada como uma ferramenta auxiliar para o cômputo de outras transformadas discretas, a DHT provou que, por si só, também apresenta utilidade em diferentes aplicações práticas, tais como processamento de imagem (PAIK; FOX, 1988), compressão de imagem (VILLASENOR, 1993; SUNDER; ESWARAN; SRIRAM, 2006), criptografia de imagem (LIU et al., 2010), interpolação de sinal (AGBINYA, 1987; MAHARANA; MEHER, 1997; TSENG; LEE, 2009), esquemas de modulação multitom (OUYANG et al., 2012; WANG;

CHANG, 2001) e marca d'água digital (MANDAL; GHOSAL, 2013). Uma coleção de algoritmos rápidos e arquiteturas de *hardware* pode ser encontrada na literatura (MEHER; SRISKANTHAN; PATRA, 2006; BRACEWELL, 1984; SORENSEN et al., 1985; GRIGORYAN, 2004; GUO, 2000; HOU, 1987a).

Versões multidimensionais para a DHT podem ser empregadas em esquemas de compressão de imagens médicas (SUNDER; ESWARAN; SRIRAAM, 2006; DULEBA, 1999; SUNDER; ESWARAN; SRIRAAM, 2005; PAPITHA; NANCY; NEDUMARAN, 2013). Conforme mencionado anteriormente, a DHT multidimensional (*mD DHT*) não apresenta a propriedade de separabilidade de núcleo (GONZALEZ; WOODS, 2001). Dessa forma, para contornar esta situação, uma transformação separável intermediária associada com algumas manipulações algébricas se faz necessária para o emprego da decomposição linha-coluna. Então, um algoritmo unidimensional para a DHT pode ser empregado a um custo aditivo extra imposto por tais manipulações algébricas.

Em (SUNDER; ESWARAN; SRIRAAM, 2006), foi proposto um esquema que explora a 3D DHT para codificação de imagens médicas, apresentando bom desempenho de compressão. Entretanto, o cômputo de baixa complexidade não é tratado em tal trabalho. *Os emergentes dispositivos biomédicos de baixa potência, tais como dispositivos biomédicos implantáveis (YAKOVLEV; KIM; POON, 2012) e miniaturizados (DESCOUR et al., 2002), apresentam severas restrições de recursos e consumo energético (GAO et al., 2011; HU; SAWAN, 2005). Embora apresente vasto leque de aplicações, métodos aproximados de baixa complexidade para a DHT são relativamente pouco explorados. Apenas em (CINTRA, 2011; BOUGUEZEL; AHMAD; SWAMY, 2013) foram apresentados métodos para derivar aproximações para a DHT, sendo este tópico ainda um campo de pesquisa em aberto. Além disso, a extensão de métodos aproximados para o caso tridimensional que leve em conta a não-separabilidade da DHT representa uma lacuna na literatura, ao nosso melhor conhecimento.*

No presente capítulo, propõe-se uma série de novas aproximações para a DHT de comprimento oito baseado na representação por racionais diádicos. As matrizes obtidas são construídas de modo a obter algoritmos livres de multiplicação e baixa complexidade computacional. O método proposto envolve dois casos: (i) abordagem de aproximações involucionais,

que consideram o mesmo algoritmo para as transformações direta e inversa; e (ii) Abordagem não-involucional, que permite diferentes transformações para a transformada inversa visando melhorar o desempenho. Também aplicamos o formalismo tensorial associado a manipulações algébricas extras para expandir as aproximações obtidas para o caso tridimensional, derivando um conjunto de aproximações para a 3D DHT. Visa-se aplicar tais aproximações tridimensionais no algoritmo de compressão de imagens baseado na 3D DHT exata proposto em (SUNDER; ESWARAN; SRIRAAM, 2006).

4.1 DHT: REVISÃO MATEMÁTICA E MÉTODOS APROXIMADOS PROPOSTOS

Na presente seção, são revisados os conceitos matemáticos necessários e são propostas matrizes aproximadas para a DHT.

4.1.1 DHT Unidimensional

A DHT de um sinal discreto $\mathbf{x} = \begin{bmatrix} x[0] & x[1] & \dots & x[N-1] \end{bmatrix}^\top$ é o sinal discreto $\mathbf{y} = \begin{bmatrix} y[0] & y[1] & \dots & y[N-1] \end{bmatrix}^\top$, definido por (BRACEWELL, 1983)

$$y[k] \triangleq \sum_{n=0}^{N-1} x[n] \cdot \text{cas} \left(\frac{2\pi kn}{N} \right), \quad (4.1)$$

$$k = 0, 1, \dots, N-1,$$

em que $\text{cas}(\cdot) \triangleq \cos(\cdot) + \sin(\cdot)$. A transformação inversa é computada de acordo com

$$x[n] = \frac{1}{N} \cdot \sum_{k=0}^{N-1} y[k] \cdot \text{cas} \left(\frac{2\pi kn}{N} \right),$$

$$n = 0, 1, \dots, N-1,$$

que é equivalente a transformada direta (4.1), exceto por um fator de escala $1/N$.

No formalismo matricial, a DHT pode ser representada pela matriz de transformação \mathbf{H}_N , cujas entradas são dadas por

$$h_N[k, n] = \text{cas} \left(\frac{2\pi kn}{N} \right), \quad k, n = 0, 1, \dots, N-1. \quad (4.2)$$

A matriz \mathbf{H}_N é ortogonal de acordo com a Definição 2.5 e simétrica (BERNSTEIN, 2009, p. 165). Logo, a seguinte propriedade é verdadeira:

$$\mathbf{H}_N^{-1} = \frac{1}{N} \cdot \mathbf{H}_N^\top = \frac{1}{N} \cdot \mathbf{H}_N.$$

Conseqüentemente, as transformadas direta e inversa são respectivamente dadas por

$$\begin{aligned} \mathbf{y} &= \mathbf{H}_N \cdot \mathbf{x}, \\ \mathbf{x} &= \frac{1}{N} \cdot \mathbf{H}_N \cdot \mathbf{y}. \end{aligned}$$

Para o comprimento $N = 8$, largamente empregado em esquemas de codificação, a matriz da DHT \mathbf{H}_8 — doravante denotada apenas por \mathbf{H} — pode ser escrita parametricamente de acordo com

$$\mathbf{H} = \begin{bmatrix} \alpha & \alpha & \alpha & \alpha & \alpha & \alpha & \alpha & \alpha \\ \alpha & \beta & \alpha & 0 & -\alpha & -\beta & -\alpha & 0 \\ \alpha & \alpha & -\alpha & -\alpha & \alpha & \alpha & -\alpha & -\alpha \\ \alpha & 0 & -\alpha & \beta & -\alpha & 0 & \alpha & -\beta \\ \alpha & -\alpha & \alpha & -\alpha & \alpha & -\alpha & \alpha & -\alpha \\ \alpha & -\beta & \alpha & 0 & -\alpha & \beta & -\alpha & 0 \\ \alpha & -\alpha & -\alpha & \alpha & \alpha & -\alpha & -\alpha & \alpha \\ \alpha & 0 & -\alpha & -\beta & -\alpha & 0 & \alpha & \beta \end{bmatrix}, \quad (4.3)$$

em que $\alpha = 1$ e $\beta = \sqrt{2}$. Tal matriz pode ser fatorada de acordo com

$$\mathbf{H} = \mathbf{A}_3 \cdot \mathbf{A}_2 \cdot \mathbf{M}(\alpha, \beta) \cdot \mathbf{A}_1 \cdot \mathbf{P}, \quad (4.4)$$

em que \mathbf{P} é uma matriz de permutação e $\mathbf{M}(\alpha, \beta)$ é uma matriz diagonal, dadas respectivamente

por

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{M}(\alpha, \beta) = \begin{bmatrix} \alpha & & & & & & & \\ & \alpha & & & & & & \\ & & \alpha & & & & & \\ & & & \alpha & & & & \\ & & & & \alpha & & & \\ & & & & & \beta & & \\ & & & & & & \alpha & \\ & & & & & & & \beta \end{bmatrix},$$

as matrizes \mathbf{A}_1 , \mathbf{A}_2 , e \mathbf{A}_3 são matrizes esparsas, respectivamente dadas por

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 1 & & & & & & \\ 1 & -1 & & & & & & \\ & & 1 & 1 & & & & \\ & & 1 & -1 & & & & \\ & & & & 1 & 1 & & \\ & & & & 1 & -1 & & \\ & & & & & & 1 & 1 \\ & & & & & & 1 & -1 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 1 & 0 & 1 & 0 & & & & \\ 0 & 1 & 0 & 1 & & & & \\ 1 & 0 & -1 & 0 & & & & \\ 0 & 1 & 0 & -1 & & & & \\ & & & & 1 & 0 & 1 & 0 \\ & & & & 0 & 1 & 0 & 0 \\ & & & & 1 & 0 & -1 & 0 \\ & & & & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{A}_3 = \begin{bmatrix} 1 & & & & 1 & & & \\ & 1 & & & & 1 & & \\ & & 1 & & & & 1 & \\ & & & 1 & & & & 1 \\ 1 & & & & -1 & & & \\ & 1 & & & & -1 & & \\ & & 1 & & & & -1 & \\ & & & 1 & & & & -1 \end{bmatrix}.$$

Note que as matrizes \mathbf{A}_1 , \mathbf{A}_2 e \mathbf{A}_3 demandam apenas operações de adição, enquanto a matriz diagonal $\mathbf{M}(\alpha, \beta)$ requer apenas multiplicações pelos parâmetros α e β . Considerando que

$\alpha = 1$ e $\beta = \sqrt{2}$, tal fatoração impõe uma complexidade aritmética de 2 multiplicações e 22 adições.

4.1.2 Aproximações para a DHT de Comprimento Oito Propostas

Nesta seção, visa-se obter aproximações de baixa complexidade para a matriz \mathbf{H} de transformação da DHT. Primeiramente, note que se uma aproximação preservar a mesma estrutura paramétrica da DHT exata explicitada em (4.3), então ela compartilhará da mesma fatoração mostrada em (4.4). Neste caso, a matriz aproximada pode ser escrita de acordo com

$$\hat{\mathbf{H}} = \begin{bmatrix} \hat{\alpha} & \hat{\alpha} & \hat{\alpha} & \hat{\alpha} & \hat{\alpha} & \hat{\alpha} & \hat{\alpha} & \hat{\alpha} \\ \hat{\alpha} & \hat{\beta} & \hat{\alpha} & 0 & -\hat{\alpha} & -\hat{\beta} & -\hat{\alpha} & 0 \\ \hat{\alpha} & \hat{\alpha} & -\hat{\alpha} & -\hat{\alpha} & \hat{\alpha} & \hat{\alpha} & -\hat{\alpha} & -\hat{\alpha} \\ \hat{\alpha} & 0 & -\hat{\alpha} & \hat{\beta} & -\hat{\alpha} & 0 & \hat{\alpha} & -\hat{\beta} \\ \hat{\alpha} & -\hat{\alpha} & \hat{\alpha} & -\hat{\alpha} & \hat{\alpha} & -\hat{\alpha} & \hat{\alpha} & -\hat{\alpha} \\ \hat{\alpha} & -\hat{\beta} & \hat{\alpha} & 0 & -\hat{\alpha} & \hat{\beta} & -\hat{\alpha} & 0 \\ \hat{\alpha} & -\hat{\alpha} & -\hat{\alpha} & \hat{\alpha} & \hat{\alpha} & -\hat{\alpha} & -\hat{\alpha} & \hat{\alpha} \\ \hat{\alpha} & 0 & -\hat{\alpha} & -\hat{\beta} & -\hat{\alpha} & 0 & \hat{\alpha} & \hat{\beta} \end{bmatrix}. \quad (4.5)$$

Entretanto, é possível mostrar que a ortogonalidade para a matriz em (4.5) é satisfeita somente se

$$\hat{\beta} = \sqrt{2} \cdot \hat{\alpha}.$$

Consequentemente, não é possível obter uma aproximação ortogonal $\hat{\mathbf{H}}$ com elementos racionais para ambos $\hat{\alpha}$ e $\hat{\beta}$ que preservem tal estrutura paramétrica. Sendo assim, focaremos no caso de transformações não ortogonais que apresentem propriedades de quase-ortogonalidade (CINTRÁ; BAYER; TABLADA, 2014).

Visa-se obter aproximações não ortogonais para a DHT livres de multiplicação e de baixo custo computacional; alto ganho de codificação C_g , dado por (2.21); baixo MSE, dado por (2.20) e baixo desvio de diagonalidade $\partial \left(\hat{\mathbf{H}}^\top \cdot \hat{\mathbf{H}} \right)$, de acordo com (2.23). Conforme dito anteriormente, separamos o problema em dois casos: (i) abordagem involucional, em que consideramos as transformadas diretas e inversas baseadas na mesma matriz, preservando a propri-

idade de involução da DHT; e (ii) abordagem não-involucional, em que a restrição da propriedade de involução é removida, visando obter uma matriz que melhor aproxime a transformada inversa segundo o critério de desvio de diagonalidade.

Preserva-se o parâmetro $\hat{\alpha} = \alpha = 1$ por ser um elemento de multiplicação trivial. Uma busca exaustiva é realizada para o parâmetro $\hat{\beta}$, variando $\hat{\beta} \in [0, 3]$ a passos diádicos de $1/8$, em que o valor exato $\beta = \sqrt{2}$ grosseiramente se encontra no meio de tal intervalo. O passo diádico é escolhido de modo a garantir que as transformadas resultantes apresentem baixa complexidade aditiva a um custo multiplicativo nulo. De fato, é possível mostrar que uma multiplicação pelos valores obtidos com tal passo do intervalo considerado pode ser implementada com não mais de 3 adições extras empregando a representação (2.22). Sendo assim, $\hat{\beta}_m = \frac{m}{8}$ para $m = 1, 2, \dots, 24$. Exclui-se o valor $\hat{\beta}_0$ uma vez que resulta em uma matriz singular. Com $\hat{\alpha} = 1$ e $\hat{\beta}_m$, é derivada uma nova matriz escrita parametricamente em função de $\hat{\beta}_m$ de acordo com

$$\hat{\mathbf{H}}(\hat{\beta}_m) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \hat{\beta}_m & 1 & 0 & -1 & -\hat{\beta}_m & -1 & 0 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 0 & -1 & \hat{\beta}_m & -1 & 0 & 1 & -\hat{\beta}_m \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -\hat{\beta}_m & 1 & 0 & -1 & \hat{\beta}_m & -1 & 0 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 0 & -1 & -\hat{\beta}_m & -1 & 0 & 1 & \hat{\beta}_m \end{bmatrix},$$

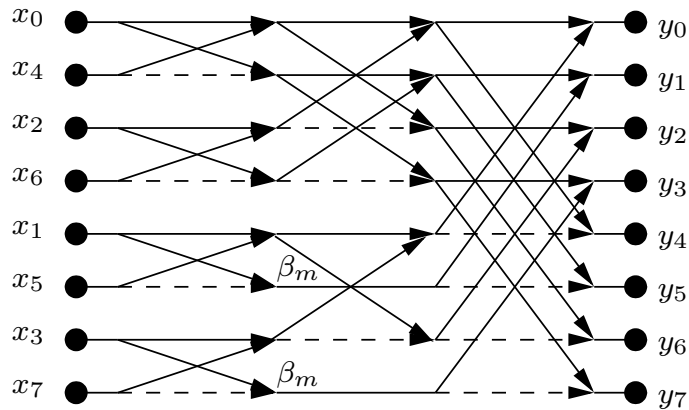
cujas fatoração é dada por

$$\hat{\mathbf{H}}(\hat{\beta}_m) = \mathbf{A}_3 \cdot \mathbf{A}_2 \cdot \mathbf{M}(1, \hat{\beta}_m) \cdot \mathbf{A}_1 \cdot \mathbf{P}.$$

O algoritmo rápido para tal fatoração é mostrado na Figura 17. A complexidade aditiva total depende da complexidade aditiva associada à multiplicação pelo elemento diádico $\hat{\beta}_m$ empregando a representação CSD. Sendo assim, tal algoritmo demanda 22 adições inerentes mais o número de adições impostas por $\hat{\beta}_m$.

Primeiramente, consideremos a (i) abordagem involucional, em que empregamos a mesma matriz para a transformada direta e inversa. Busca-se pelas aproximações ótimas de acordo com

Figura 17 – Algoritmo rápido para a DHT de comprimento oito.



Fonte: O autor

cada uma das métricas: desvio de diagonalidade $\partial [\hat{\mathbf{H}}(\hat{\beta}_m) \cdot \hat{\mathbf{H}}(\hat{\beta}_m)]$, MSE, ganho de codificação e a complexidade computacional. Em seguida, para a (ii) abordagem não-involucional, realiza-se a mesma para ambas as transformadas direta e inversa. Neste caso, permite-se que as transformadas direta e inversa sejam diferentes. Para cada matriz aproximada $\hat{\mathbf{H}}(\hat{\beta}_m)$ obtida, busca-se por uma matriz aproximada $\hat{\mathbf{H}}(\hat{\beta}_q)$ que minimize o desvio de diagonalidade $\partial [\hat{\mathbf{H}}(\hat{\beta}_m) \cdot \hat{\mathbf{H}}(\hat{\beta}_q)]$.

Como resultado, a matriz $\hat{\mathbf{H}}(\hat{\beta}_{11})$ apresenta o menor valor de MSE; a matriz $\hat{\mathbf{H}}(\hat{\beta}_{12})$ possui o maior ganho de codificação; e a matriz $\hat{\mathbf{H}}(\hat{\beta}_8)$ mostra o menor custo computacional. Para a abordagem não-involucional, encontramos que o desvio de diagonalidade reduz quando a matriz $\hat{\mathbf{H}}(\hat{\beta}_{12})$ é tomada como a inversa aproximada de $\hat{\mathbf{H}}(\hat{\beta}_{11})$ e vice-versa. Ainda mais, a matriz $\hat{\mathbf{H}}(\hat{\beta}_8)$ apresenta como matriz inversa *exata* a matriz $\hat{\mathbf{H}}(\hat{\beta}_{16})$, ou seja, $\partial [\hat{\mathbf{H}}(\hat{\beta}_8) \cdot \hat{\mathbf{H}}(\hat{\beta}_{16})] = 0$. A Tabela 12 mostra a representação numérica e a complexidade associada para cada parâmetro β_m das matrizes derivadas. As métricas obtidas para cada aproximação são resumidas na Tabela 13.

4.2 DHT TRIDIMENSIONAL: REVISÃO, DEFINIÇÃO MATEMÁTICA E FORMALIZAÇÃO PARA APROXIMAÇÕES

Nesta seção, abordamos as definições matemáticas para a 3D DHT exata e aproximada. Desenvolve-se a definição da 3D DHT baseada no formalismo tensorial exposto na Seção 4.2.1. Na Seção 4.2.2, propõe-se uma expressão algébrica geral para a 3D DHT aproximada baseada

Tabela 12 – Representação e complexidade dos parâmetros β_m para cada matriz aproximada para a DHT

m	β_m	Representação CSD	Complexidade	
			Adições	Deslocamentos de bits
8	1	1	0	0
11	11/8	$1 + \frac{1}{4} + \frac{1}{8}$	2	2
12	3/2	$1 + \frac{1}{2}$	1	1
16	2	2	0	1

Fonte: o autor

Tabela 13 – Métricas e complexidades computacionais para as matrizes aproximadas obtidas para a DHT

Direta	Inversa	Métricas			Complex. $\hat{\mathbf{H}}(\hat{\beta}_m)$	
		$\partial [\hat{\mathbf{H}}(\hat{\beta}_m) \cdot \hat{\mathbf{H}}(\hat{\beta}_q)]$	C_g (dB)	MSE	Ad.	Desloc.
$\hat{\mathbf{H}}(1)$	$\hat{\mathbf{H}}(1)$	$1.94 \cdot 10^{-2}$	7.418	$3.182 \cdot 10^{-2}$	22	0
$\hat{\mathbf{H}}(\frac{11}{8})$	$\hat{\mathbf{H}}(\frac{11}{8})$	$1.92 \cdot 10^{-4}$	7.818	$2.85 \cdot 10^{-4}$	24	2
$\hat{\mathbf{H}}(\frac{3}{2})$	$\hat{\mathbf{H}}(\frac{3}{2})$	$9.16 \cdot 10^{-4}$	7.830	$13.65 \cdot 10^{-4}$	23	1
$\hat{\mathbf{H}}(1)$	$\hat{\mathbf{H}}(2)$	0	7.418	$3.182 \cdot 10^{-2}$	22	0
$\hat{\mathbf{H}}(\frac{11}{8})$	$\hat{\mathbf{H}}(\frac{3}{2})$	$6.01 \cdot 10^{-5}$	7.818	$2.852 \cdot 10^{-4}$	24	2
$\hat{\mathbf{H}}(\frac{3}{2})$	$\hat{\mathbf{H}}(\frac{11}{8})$	$6.01 \cdot 10^{-5}$	7.830	$1.365 \cdot 10^{-3}$	23	1
$\hat{\mathbf{H}}(2)$	$\hat{\mathbf{H}}(1)$	0	7.506	$6.365 \cdot 10^{-2}$	22	1

Fonte: o autor

em matrizes aproximadas. As aproximações para a DHT de comprimento oito derivadas na Seção 4.1.2 foram, então, empregadas para derivar novas aproximações para a 3D DHT de dimensões $8 \times 8 \times 8$. A complexidade aritmética é discutida na Seção 4.2.3.

4.2.1 DHT Tridimensional Exata

A 3D DHT de um tensor de terceira ordem $\mathcal{X} \in \mathbb{F}^{N_1 \times N_2 \times N_3}$, cujas entradas são $x[n_1, n_2, n_3]$, para $n_i = 0, 1, \dots, N_i - 1$ e $i = 1, 2, 3$, é definida como o tensor de terceira ordem $\mathcal{Y} \in \mathbb{F}^{N_1 \times N_2 \times N_3}$, cujas entradas são dadas por

$$y[k_1, k_2, k_3] \triangleq \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \sum_{n_3=0}^{N_3-1} x[n_1, n_2, n_3] \cdot \text{cas} \left(\frac{2\pi n_1 k_1}{N_1} + \frac{2\pi n_2 k_2}{N_2} + \frac{2\pi n_3 k_3}{N_3} \right), \quad (4.6)$$

$$k_i = 0, 1, \dots, N_i - 1, \quad i = 1, 2, 3.$$

De forma similar ao caso 1D, a 3D DHT inversa é idêntica à direta, exceto por um fator de escalonamento, de acordo com

$$x[n_1, n_2, n_3] = \frac{1}{N_1 N_2 N_3} \cdot \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \sum_{k_3=0}^{N_3-1} y[k_1, k_2, k_3] \cdot \text{cas} \left(\frac{2\pi n_1 k_1}{N_1} + \frac{2\pi n_2 k_2}{N_2} + \frac{2\pi n_3 k_3}{N_3} \right),$$

$$n_i = 0, 1, \dots, N_i - 1, \quad i = 1, 2, 3.$$
(4.7)

Diferentemente do caso da 3D DCT, a 3D DHT não apresenta a propriedade de separabilidade de acordo com a Definição 2.2. Para contornar este comportamento da 3D DHT, define-se a 3D DHT *especial* (3D SDHT), definida com núcleo separável, dada por um tensor $\mathcal{Y}_s \in \mathbb{F}^{N_1 \times N_2 \times N_3}$, cujas entradas são definidas como

$$y_s[k_1, k_2, k_3] \triangleq \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \sum_{n_3=0}^{N_3-1} x[n_1, n_2, n_3] \cdot \text{cas} \left(\frac{2\pi n_1 k_1}{N_1} \right) \cdot \text{cas} \left(\frac{2\pi n_2 k_2}{N_2} \right) \cdot \text{cas} \left(\frac{2\pi n_3 k_3}{N_3} \right)$$

$$= \sum_{n_1=0}^{N_1-1} x[n_1, n_2, n_3] \cdot \text{cas} \left(\frac{2\pi n_1 k_1}{N_1} \right) \cdot \sum_{n_2=0}^{N_2-1} \text{cas} \left(\frac{2\pi n_2 k_2}{N_2} \right) \cdot \sum_{n_3=0}^{N_3-1} \text{cas} \left(\frac{2\pi n_3 k_3}{N_3} \right),$$

$$k_i = 0, 1, \dots, N_i - 1, \quad i = 1, 2, 3.$$
(4.8)

Dessa forma, a 3D SDHT pode ser expressa em termos de produtos em modo i -ésimo (3.6) pelas matrizes da DHT em cada dimensão (4.2), de acordo com

$$\mathcal{Y}_s = \mathcal{X} \times_1 \mathbf{H}_{N_1} \times_2 \mathbf{H}_{N_2} \times_3 \mathbf{H}_{N_3}. \quad (4.9)$$

Levando em consideração que (HAO; BRACEWELL, 1987)

$$\begin{aligned} \text{cas}(\alpha + \beta + \gamma) &= \frac{1}{2} [\text{cas}(-\alpha) \cdot \text{cas}(\beta) \cdot \text{cas}(\gamma) \\ &\quad + \text{cas}(\alpha) \cdot \text{cas}(-\beta) \cdot \text{cas}(\gamma) \\ &\quad + \text{cas}(\alpha) \cdot \text{cas}(\beta) \cdot \text{cas}(-\gamma) \\ &\quad - \text{cas}(-\alpha) \cdot \text{cas}(-\beta) \cdot \text{cas}(-\gamma)], \end{aligned}$$

o tensor \mathcal{Y} da 3D DHT pode ser expresso em função de versões permutadas do tensor \mathcal{Y}_s da 3D SDHT, conforme o que se segue:

$$\mathcal{Y} = \frac{1}{2} (\mathcal{Y}_s^{(1)} + \mathcal{Y}_s^{(2)} + \mathcal{Y}_s^{(3)} - \mathcal{Y}_s^{(4)}),$$

em que as entradas de $\mathcal{Y}_s^{(1)}$, $\mathcal{Y}_s^{(2)}$, $\mathcal{Y}_s^{(3)}$ e $\mathcal{Y}_s^{(4)}$ são respectivamente dadas por

$$\begin{aligned}
y_s^{(1)}[k_1, k_2, k_3] &= y_s \left[((N_1 - k_1))_{N_1}, k_2, k_3 \right], \\
y_s^{(2)}[k_1, k_2, k_3] &= y_s \left[k_1, ((N_2 - k_2))_{N_2}, k_3 \right], \\
y_s^{(3)}[k_1, k_2, k_3] &= y_s \left[k_1, k_2, ((N_3 - k_3))_{N_3} \right], \\
y_s^{(4)}[k_1, k_2, k_3] &= y_s \left[((N_1 - k_1))_{N_1}, ((N_2 - k_2))_{N_2}, ((N_3 - k_3))_{N_3} \right], \\
k_i &= 0, 1, \dots, N_i - 1, \quad i = 1, 2, 3,
\end{aligned} \tag{4.10}$$

e $((k))_N$ indica $(k \bmod N)$ (OPPENHEIM; SCHAFER, 2010, p. 142). A transformada inversa é computada por um procedimento similar que inclui o fator de escala. Então, tem-se que

$$\mathcal{X}_s = \frac{1}{2N_1N_2N_3} \cdot (\mathcal{Y} \times_1 \mathbf{H}_{N_1} \times_2 \mathbf{H}_{N_2} \times_3 \mathbf{H}_{N_3})$$

e

$$\mathcal{X} = \mathcal{X}_s^{(1)} + \mathcal{X}_s^{(2)} + \mathcal{X}_s^{(3)} - \mathcal{X}_s^{(4)}.$$

4.2.2 Formalização Matemática para o Caso Tridimensional

O objetivo desta seção é derivar uma abordagem geral para aproximar a 3D DHT, definida em (4.6) e (4.7). Visa-se propor um método baseado nas matrizes aproximadas para a DHT propostas na Seção 4.1.2. As motivações para tal abordagem são: (i) a 3D DHT aproximada estaria matematicamente relacionada com a versão unidimensional de forma direta; e (ii) um algoritmo rápido multidimensional baseado no método linha-coluna pode ser proposto empregando um algoritmo rápido para a DHT unidimensional aproximada. Nesse caso, ao invés de aproximar diretamente (4.6), busca-se primeiramente aproximar a SDHT em (4.8) de acordo com o formalismo tensorial descrito em (4.9) e (4.10). Seja $\mathcal{X} \in \mathbb{F}^{N_1 \times N_2 \times N_3}$ um tensor de terceira ordem arbitrário e $\hat{\mathbf{H}}_{N_1}$, $\hat{\mathbf{H}}_{N_2}$, $\hat{\mathbf{H}}_{N_3}$ aproximações arbitrárias para a matriz da DHT. De maneira análoga ao caso da 3D DCT aproximada descrita na Seção 3.2, define-se a 3D SDHT aproximada substituindo as matrizes exatas em (4.9) por respectivas matrizes aproximadas em cada dimensão de acordo com

$$\mathcal{Y}_s \triangleq \mathcal{X} \times_1 \hat{\mathbf{H}}_{N_1} \times_2 \hat{\mathbf{H}}_{N_2} \times_3 \hat{\mathbf{H}}_{N_3}, \tag{4.11}$$

em que \mathcal{Y}_s é tensor de saída da 3D SDHT aproximada. Então, define-se a 3D DHT aproximada a partir da 3D SDHT aproximada de forma idêntica ao caso exato. Dessa forma, tem-se que

$$\mathcal{Y} \triangleq \frac{1}{2} \cdot (\mathcal{Y}_s^{(1)} + \mathcal{Y}_s^{(2)} + \mathcal{Y}_s^{(3)} - \mathcal{Y}_s^{(4)}). \quad (4.12)$$

A transformada inversa é derivada de forma similar. Inicialmente, computa-se

$$\begin{aligned} \mathcal{X}_s &= \mathcal{Y} \times_1 \hat{\mathbf{H}}_{N_1}^{-1} \times_2 \hat{\mathbf{H}}_{N_2}^{-1} \times_3 \hat{\mathbf{H}}_{N_3}^{-1} \\ &= \mathcal{Y} \times_1 \left(\hat{\mathbf{H}}_{N_1} \cdot \mathbf{D}_{N_1} \right) \times_2 \left(\hat{\mathbf{H}}_{N_2} \cdot \mathbf{D}_{N_2} \right) \times_3 \left(\hat{\mathbf{H}}_{N_3} \cdot \mathbf{D}_{N_3} \right) \\ &= \mathcal{Y} \times_1 \mathbf{D}_{N_1} \times_2 \mathbf{D}_{N_2} \times_3 \mathbf{D}_{N_3} \times_1 \hat{\mathbf{H}}_{N_1} \times_2 \hat{\mathbf{H}}_{N_2} \times_3 \hat{\mathbf{H}}_{N_3}, \end{aligned} \quad (4.13)$$

em que \mathbf{D}_N é a matriz diagonal dada em (2.19). Em (4.13), aplicam-se as Proposições 3.1 e 3.2. Conforme descrito no Capítulo 3, a matriz diagonal pode ser incorporada na etapa de desquantização, eliminando completamente a sobrecarga computacional imposta por tal matriz (BRITANAK; YIP; RAO, 2007; SUNDER; ESWARAN; SRIRAAM, 2006; COUTINHO; CINTRA; BAYER, 2017). Finalmente, obtém-se

$$\mathcal{X} = \frac{1}{2} \cdot (\mathcal{X}_s^{(1)} + \mathcal{X}_s^{(2)} + \mathcal{X}_s^{(3)} - \mathcal{X}_s^{(4)}).$$

Para o caso particular em que $N_1 = N_2 = N_3 \triangleq N$, uma mesma matriz aproximada $\hat{\mathbf{H}}_N$ pode ser empregada para computar uma 3D DHT aproximada fazendo $\hat{\mathbf{H}}_{N_1} = \hat{\mathbf{H}}_{N_2} = \hat{\mathbf{H}}_{N_3} \triangleq \hat{\mathbf{H}}_N$. Então, tem-se que

$$\mathcal{Y}_s = \mathcal{X} \times_1 \hat{\mathbf{H}}_N \times_2 \hat{\mathbf{H}}_N \times_3 \hat{\mathbf{H}}_N. \quad (4.14)$$

Doravante, focaremos em tal particular caso com $N = 8$. Consideraremos as matrizes derivadas na Seção 4.1.2 em (4.14) para obter um conjunto de aproximações para a 3D DHT.

4.2.3 Análise de Complexidade

A análise de complexidade para a 3D DHT é análoga ao caso da 3D DCT, conforme descrito na Seção 3.2.2, tendo como principal diferença as operações adicionais impostas pela não separabilidade da 3D DHT.

Seja $A_{1D}(\mathbf{H}_N)$ a complexidade aritmética da transformação unidimensional (2.18) representada por \mathbf{H}_N . Então, a complexidade aritmética para a 3D SDHT pode ser obtida em

função da complexidade do algoritmo unidimensional de acordo com

$$C_{3D}(\mathbf{H}_N) = 3N^2 \cdot C_{1D}(\mathbf{H}_N). \quad (4.15)$$

Para computar a 3D DHT a partir da 3D SDHT empregando as expressões (4.12), uma sobrecarga de $3N^3$ adições é necessária. Dessa forma, a complexidade aditiva total para a 3D DHT é dada por

$$A_{3D}(\mathbf{H}_N) = 3N^2 \cdot A_{1D}(\mathbf{H}_N) + 3N^3. \quad (4.16)$$

As matrizes obtidas na Seção 4.1.2 foram aplicadas em (4.14). Um conjunto de aproximações para a 3D DHT é obtido baseado no formalismo descrito anteriormente. A Tabela 14 apresenta a complexidade aritmética total para cada aproximação de dimensões $8 \times 8 \times 8$ proposta. Para levantar as complexidades das aproximações para a 3D DHT, considera-se as complexidades dos algoritmos unidimensionais nas expressões (4.15) e (4.16). Para fins de comparação, está inclusa a complexidade aritmética da 3D DHT exata de acordo com a decomposição linha coluna e também a complexidade do algoritmo de base mista proposto em (BOUSSAKTA; ALSHIBAMI; AZIZ, 2001).

A aproximação proposta baseada na matriz $\hat{\mathbf{H}}(1)$ apresenta a mesma complexidade aditiva do método exato. Já os demais métodos propostos apresentam custo aditivo ligeiramente maior do que o método exato. Entretanto, todos os métodos propostos apresentam custo computacional nulo, o que é uma vantagem significativa uma vez que operações de multiplicação demandam mais recursos de energia e área de *hardware* (BLAHUT, 2010; OLIVEIRA et al., 2017; LIANG; TRAN, 2001).

4.3 APLICAÇÃO EM COMPRESSÃO DE IMAGENS MÉDICAS

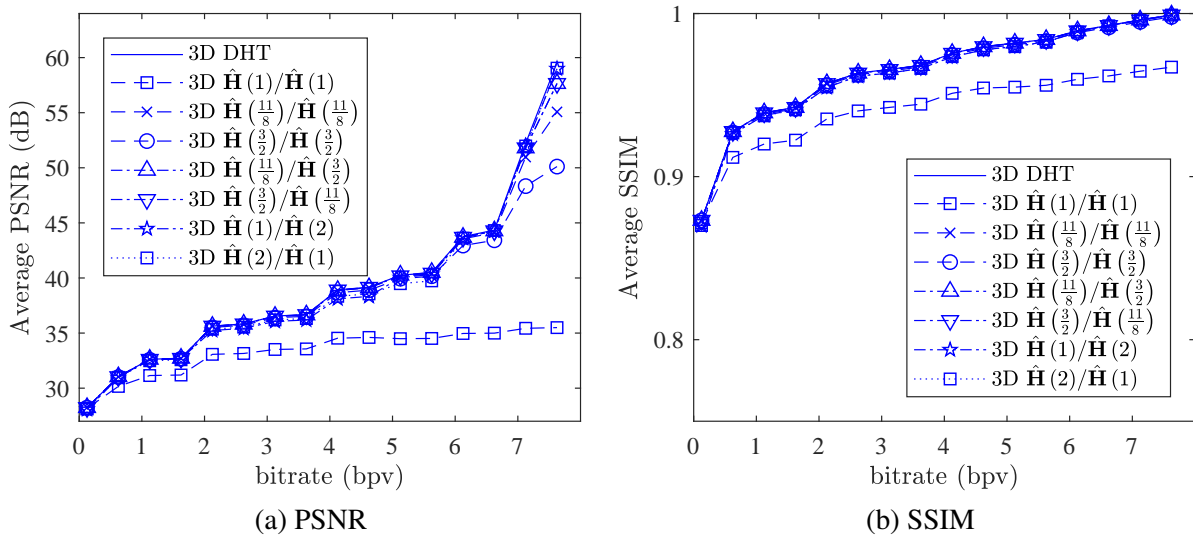
As aproximações para a 3D DHT propostas na Seção 4.2 foram consideradas no esquema de codificação proposto em (SUNDER; ESWARAN; SRIRAAM, 2006). Foram selecionados dez angiografias por raios-X (XA) e dez ressonâncias magnéticas (MR) disponíveis em (Pixmeo SARL Company, ; Rubo Medical Imaging, 2018; Zenodo.org, ; BARRÉ,) com resoluções variadas de 256×256 ou 512×512 e variado número de quadros ou *slices*.

Tabela 14 – Complexidade computacional para os algoritmos para a 3D DHT e versões aproximadas de dimensões $8 \times 8 \times 8$

Método	Complexidade 3D		
	Multiplicações	Adições	Desloc. de <i>bits</i>
3D DHT (decomposição linha-coluna)	384	5760	0
3D DHT base mista	384	5760	0
Transformada 3D proposta baseada em $\hat{\mathbf{H}}(1)$	0	5760	0
Transformada 3D proposta baseada em $\hat{\mathbf{H}}(\frac{11}{8})$	0	6528	768
Transformada 3D proposta baseada em $\hat{\mathbf{H}}(\frac{3}{2})$	0	6144	384
Transformada 3D proposta baseada em $\hat{\mathbf{H}}(2)$	0	5760	384

Fonte: o autor

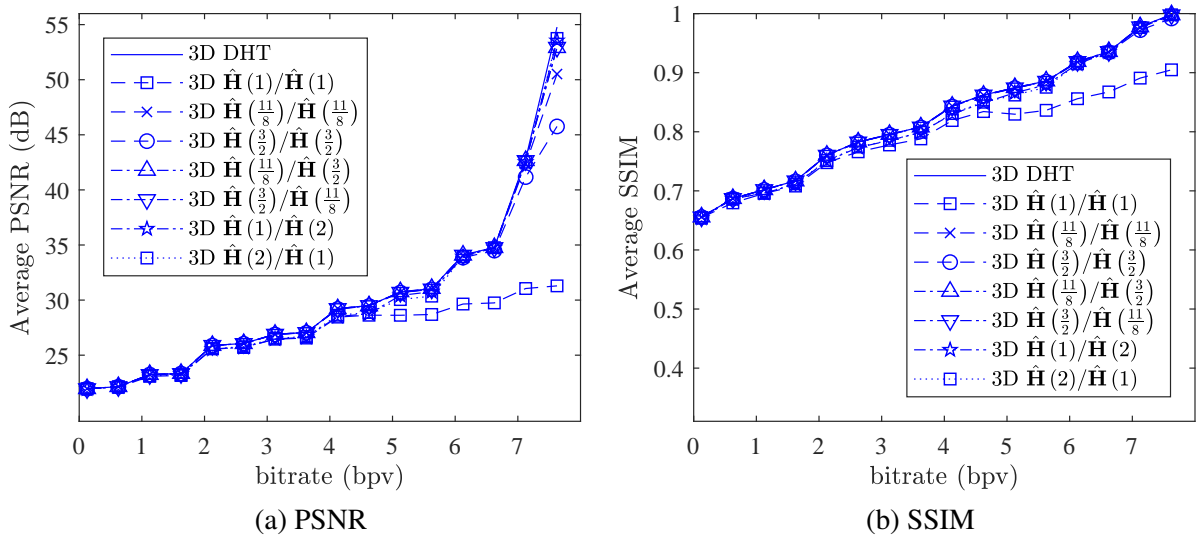
Figura 18 – PSNR e SSIM médios versus *bitrate* para a 3D DHT exata e cada método aproximado proposto em compressão de arranjos XA.



Fonte: o autor

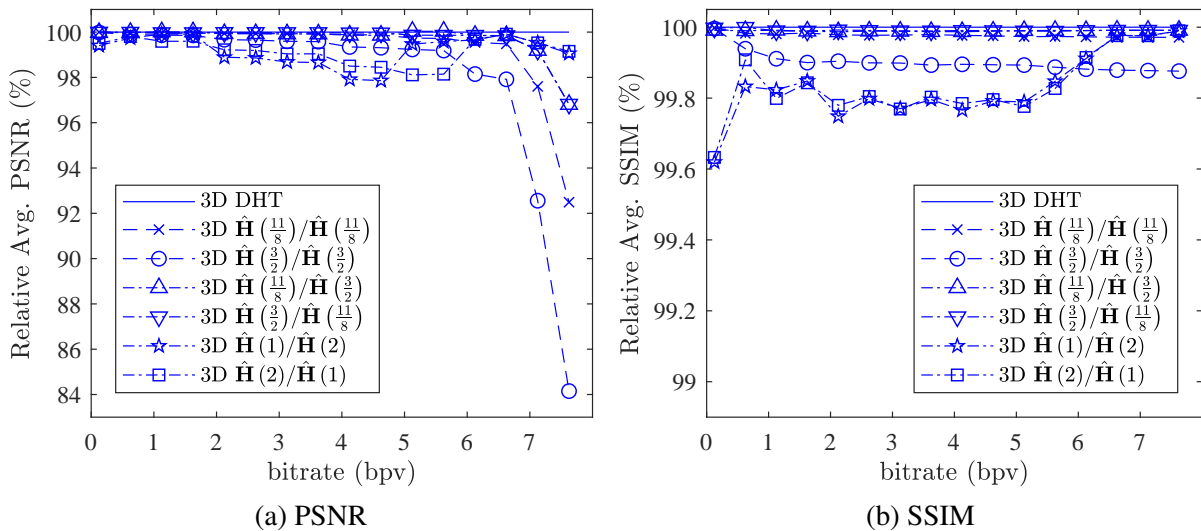
Cada XA ou MR pode ser representada por um arranjo tridimensional, isto é, um tensor de terceira ordem conforme descrito na Seção 3.1.2. Cada tensor de terceira ordem é dividido em blocos menores de dimensões $8 \times 8 \times 8$ e são submetidos às transformadas 3D conforme definido em (4.11) e (4.12). O bloco transformado é, então, aplicado ao esquema de escaneamento *zig-zag* definido em (SUNDEK; ESWARAN; SRIRAAM, 2006). Subsequentemente, os dados são comprimidos a uma taxa de bits (*bitrate*) fixa (BHASKARAN; KONSTANTINIDES, 1997; BOUGUEZEL; AHMAD; SWAMY, 2008) preservando apenas um número constante dos primeiros coeficientes do vetor *zig-zag*, que geralmente apresentam a maior parte da energia do sinal.

Figura 19 – PSNR e SSIM médios versus *bitrate* para a 3D DHT exata e cada método aproximado proposto em compressão de arranjos MR.



Fonte: o autor

Figura 20 – PSNR e SSIM relativos versus *bitrate* dos métodos propostos comparados com a 3D DHT exata em compressão de arranjos XA.



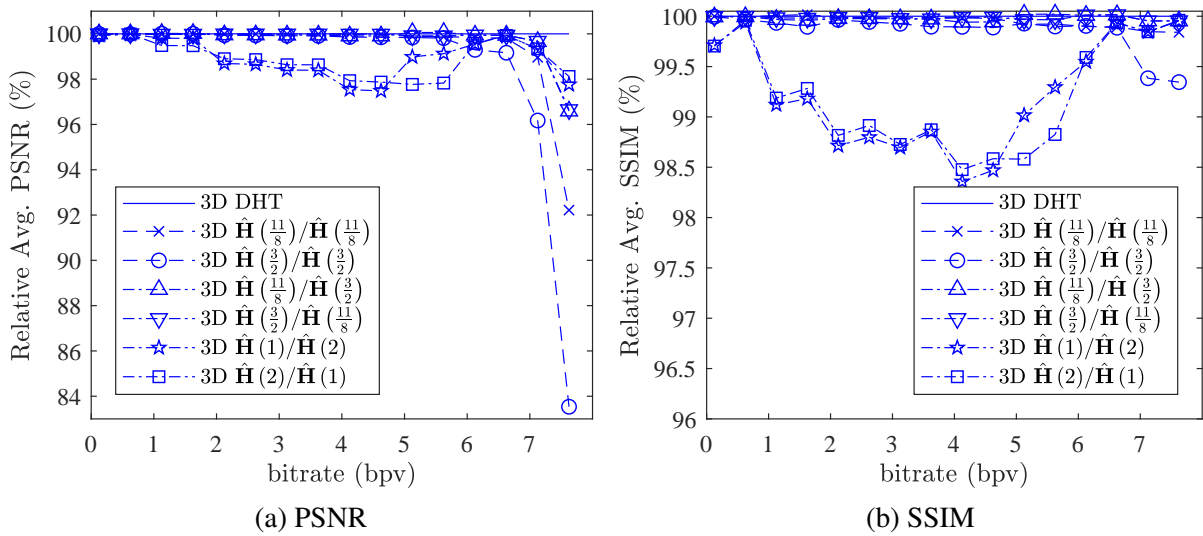
Fonte: o autor

O *bitrate* é dado em *bit per voxel* (bpv) e calculado de acordo com (SUNDER; ESWARAN; SRIRAAM, 2006):

$$\text{bitrate} = \frac{T_b}{T_v},$$

em que T_b é o tamanho da imagem comprimida em total de *bits* e T_v é o total de *voxels* da imagem. Para blocos de dimensões $8 \times 8 \times 8$, $T_v = 512$. Variando o total de coeficientes

Figura 21 – PSNR e SSIM relativos versus *bitrate* dos métodos propostos comparados com a 3D DHT exata em compressão de arranjos MR.



Fonte: o autor

preservados do vetor *zig-zag*, varia-se o *bitrate* de cada imagem no intervalo $[0.125, 7.125]$ com passos de 0.5. Como figuras de mérito para avaliação de desempenho, considera-se o PSNR e o SSIM, definidos na Seção 2.3.

A média entre cada métrica obtida de cada arranjo é levada em conta e ponderada de acordo com o número de blocos de cada arranjo. Os valores médios de PSNR e SSIM em função do *bitrate* são mostrados para XA na Figura 18 e para MR na Figura 19. Denota-se o método codificador empregando $\hat{\mathbf{H}}(\hat{\beta}_m)$ e o decodificador empregando $\hat{\mathbf{H}}(\hat{\beta}_q)$ simplesmente por $\hat{\mathbf{H}}(\hat{\beta}_m)/\hat{\mathbf{H}}(\hat{\beta}_q)$.

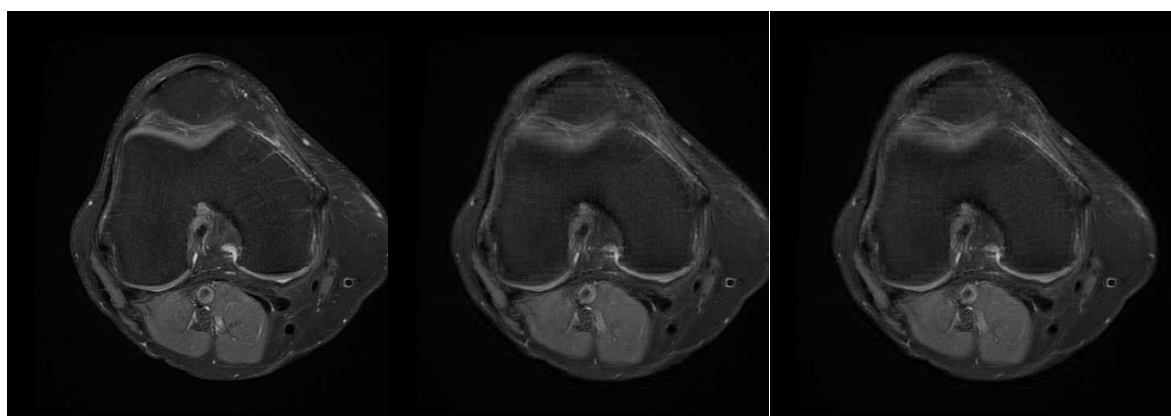
Exceto pelo método $\hat{\mathbf{H}}(1)/\hat{\mathbf{H}}(1)$, os demais métodos aproximados propostos mostram desempenho muito próximo do método usual baseado na 3D DHT exata, com as curvas de PSNR e SSIM apresentando praticamente o mesmo comportamento. De fato, o método $\hat{\mathbf{H}}(1)/\hat{\mathbf{H}}(1)$ tem desempenho substancialmente inferior por possuir o maior desvio de diagonalidade, como mostrado na Tabela 13. Sendo assim, exclui-se tal método da análise a seguir.

Para uma melhor visualização, são mostrado nas Figura 20 e Figura 21 os valores relativos de PSNR e SSIM comparados ao método baseado na 3D DHT exata. Para valores baixos de *bitrates* (< 2 bpv), o desempenho geral é superior a 99%. Especificamente para imagens XA, o desempenho dos métodos propostos é superior a 99.5% em termos de SSIM. As simulações

sugerem, que todos os métodos — incluindo o método exato — apresentam melhor desempenho em codificação de imagens XA do que imagens MR. De forma geral, todos os métodos apresentam desempenho superior a 97% para a maioria dos valores de *bitrate*, a um custo consideravelmente menor.

Uma comparação qualitativa é apresentada na Figura 22 e na Figura 23 para um *slice* de MR e XA, respectivamente, comprimidos a 4 bpv. As imagens comprimidas são muito próximas da imagem original não comprimida e os métodos aproximados apresentam praticamente o mesmo desempenho do método baseado na 3D DHT. *Tais resultados sugerem que os métodos aproximados para a 3D DHT são apropriados para sistemas de codificação de imagem médica tal qual o método em (SUNDER; ESWARAN; SRIRAAM, 2006) em contextos de restritos recursos de hardware e potência.*

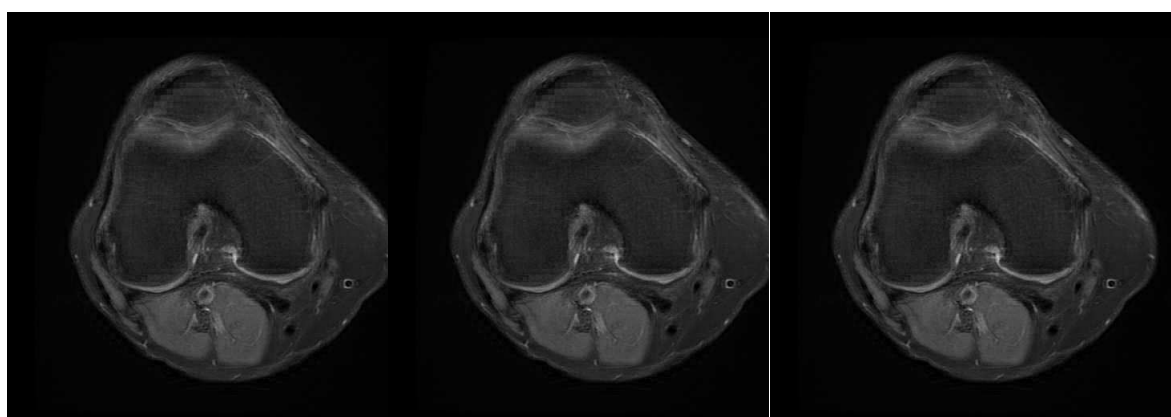
Figura 22 – Avaliação qualitativa para um slice de MR comprimido a 4 bpv empregando a 3D DHT exata e várias versões aproximadas propostas.



(a) *Slice* MR original

(b) 3D DHT (PSNR=33.07, SSIM=0.919)

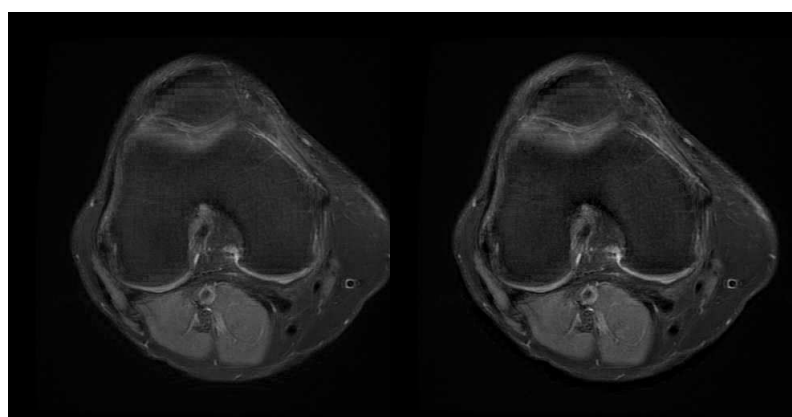
(c) 3D $\hat{H}(\frac{11}{8}) / \hat{H}(\frac{11}{8})$
(PSNR=33.10, SSIM=0.920)



(d) 3D $\hat{H}(\frac{3}{2}) / \hat{H}(\frac{3}{2})$
(PSNR=32.96, SSIM=0.917)

(e) 3D $\hat{H}(\frac{11}{8}) / \hat{H}(\frac{3}{2})$
(PSNR=33.12, SSIM=0.920)

(f) 3D $\hat{H}(\frac{3}{2}) / \hat{H}(\frac{11}{8})$
(PSNR=32.95, SSIM=0.917)

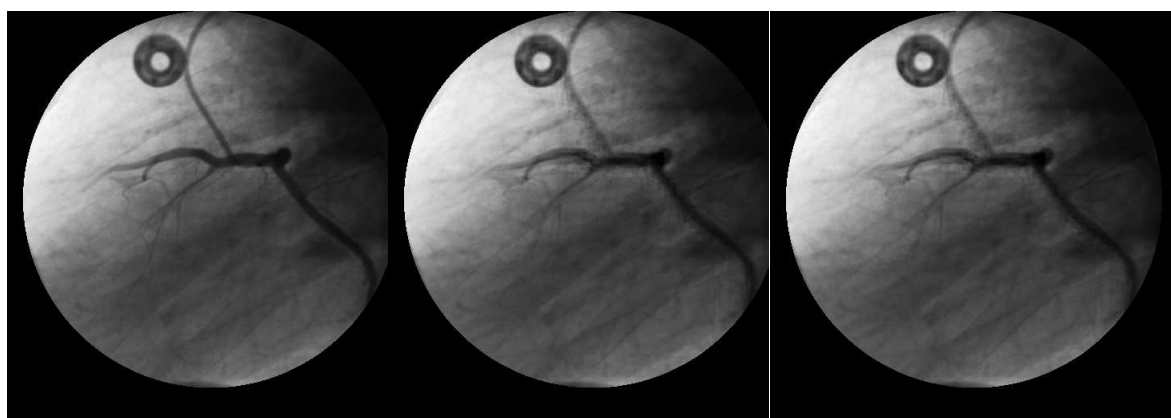


(g) 3D $\hat{H}(1) / \hat{H}(2)$
(PSNR=33.36, SSIM=0.925)

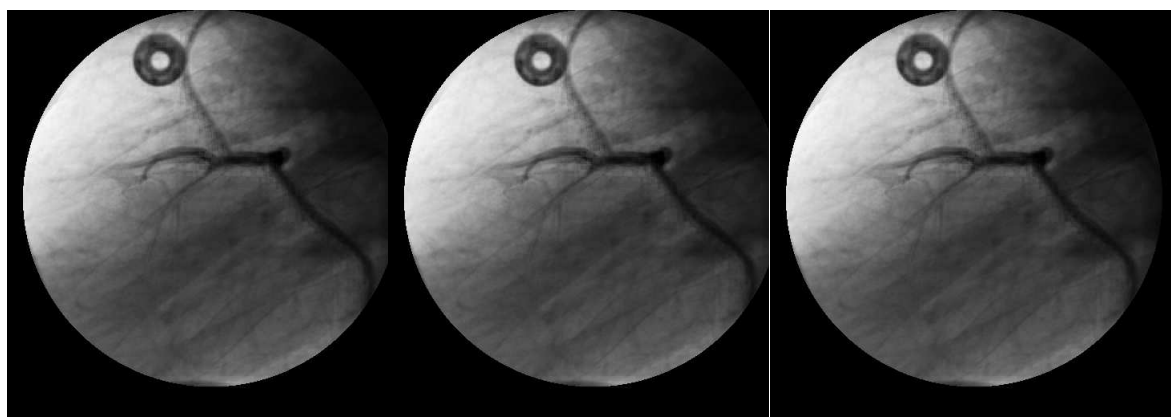
(h) 3D $\hat{H}(2) / \hat{H}(1)$
(PSNR=31.75, SSIM=0.895)

Fonte: o autor

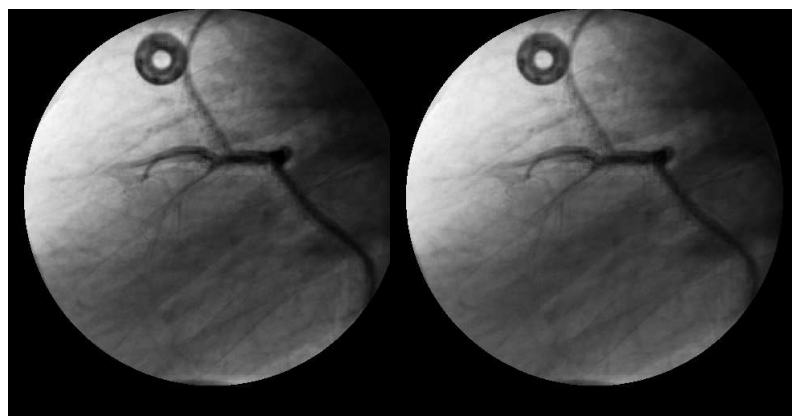
Figura 23 – Avaliação qualitativa para um slice de XA comprimido a 4 bpv empregando a 3D DHT exata e várias versões aproximadas propostas.



(a) *Slice* XA original (b) 3D DHT (PSNR=37.36, SSIM=0.970) (c) 3D $\hat{H}(\frac{11}{8}) / \hat{H}(\frac{11}{8})$ (PSNR=37.35, SSIM=0.969)



(d) 3D $\hat{H}(\frac{3}{2}) / \hat{H}(\frac{3}{2})$ (PSNR=37.100, SSIM=0.969) (e) 3D $\hat{H}(\frac{11}{8}) / \hat{H}(\frac{3}{2})$ (PSNR=37.209, SSIM=0.969) (f) 3D $\hat{H}(\frac{3}{2}) / \hat{H}(\frac{11}{8})$ (PSNR=37.424, SSIM=0.970)



(g) 3D $\hat{H}(1) / \hat{H}(2)$ (PSNR=36.110, SSIM=0.963) (h) 3D $\hat{H}(2) / \hat{H}(1)$ (PSNR=37.86, SSIM=0.972)

Fonte: o autor

5 CONCLUSÕES

No Capítulo 2, foi introduzido um conjunto de aproximações podadas de baixa complexidade para a 2D DTT. A complexidade aritmética dos métodos propostos foram analisadas e mostraram uma redução significativa no total de operações quando comparadas com a DTT exata e aproximada não podada. Simulações do tipo JPEG foram realizadas e demonstraram a eficiência dos métodos propostos para compressão de imagem estática, sugerindo um *trade-off* favorável entre complexidade e desempenho. Arquiteturas de *hardware* foram realizadas em FPGA e o consumo de recursos pelos métodos propostos apresentou uma economia considerável ao custo de uma aceitável redução de desempenho em compressão.

No Capítulo 3, foi proposta uma formulação algébrica para aproximações lineares para a 3D DCT empregando o formalismo tensorial. Diversas aproximações de baixa complexidade para a 3D DCT foram sugeridas baseadas em matrizes aproximadas do estado-da-arte. O conceito foi generalizado matematicamente para o caso multidimensional. Uma análise generalizada de complexidade foi realizada. Expressões matemáticas para a complexidade multidimensional foram derivadas. Tais expressões foram consideradas para levantar as complexidades computacionais dos métodos aproximados propostos para a 3D DCT. Uma análise de *trade-off* foi realizada considerando complexidade computacional e ganho de codificação, em que os resultados mostraram que as aproximações propostas para a 3D DCT são adequadas na maioria dos cenários.

As aproximações propostas para a 3D DCT foram empregadas em um esquema de codificação entre quadros de vídeo. Neste contexto, foi proposto um procedimento para modificação de um dado volume de quantização visando explorar as propriedades de baixa complexidade dos métodos propostos. Os resultados das simulações mostraram que os métodos propostos apresentam um desempenho competitivo quando comparado à 3D DCT exata. Também foi simulado um sistema de rastreamento visual considerando o método aproximado 3D MRDCT. Tal método é originalmente baseado na 3D DCT exata. Os resultados mostraram que o método modificado de baixa complexidade apresenta desempenho similar ao esquema original. De forma geral,

as aproximações propostas para a 3D DCT podem ser efetivamente consideradas como ferramentas de baixa complexidade para as emergentes tecnologias baseadas na 3D DCT exata que apresentem restrições de recursos de *hardware*, tal qual o sistema SoftCast de *broadcasting* (JAKUBCZAK; RAHUL; KATABI, 2010; JAKUBCZAK; KATABI, 2010; JAKUBCZAK; KATABI, 2011), codificadores de vídeo 3D baseados em quantização (MULLA et al., 2014; SAPONARA, 2012b; SAWANT; ADJEROH, 2011; BOZINOVIĆ; KONRAD, 2003; LAI; GUAN, 2002; CHAN; SIU, 1997; LEE; CHAN; ADJEROH, 1997) e rastreadores visuais baseados em transformadas de bloco (LI et al., 2013).

Também foram propostas, no Capítulo 4, aproximações para a matriz da DHT baseadas em uma busca exaustiva paramétrica. Um conjunto de matrizes aproximadas quase-ortogonais foi apresentado. Tais matrizes foram otimizadas de acordo com medidas de desempenho, complexidade e quase-ortogonalidade. Também foi proposta uma formalização matemática para derivar aproximações para a 3D DHT que considera as características de não separabilidade da 3D DHT exata e que se baseia em matrizes aproximadas para a DHT. Foi proposta uma coleção de aproximações de baixa complexidade para a 3D DHT de dimensão $8 \times 8 \times 8$ empregando as matrizes derivadas e o formalismo proposto. Os métodos 3D derivados foram empregados em um esquema de compressão de imagens médicas, originalmente baseado na 3D DHT exata. Os métodos propostos apresentaram um desempenho praticamente igual ao método exato a um custo computacional consideravelmente reduzido. Sendo assim, os métodos propostos são ferramentas realísticas para aplicações em sistemas de compressão, transmissão e armazenamento de imagens médicas com rigorosas restrições de *hardware*.

5.1 CONTRIBUIÇÕES

As contribuições derivadas do presente trabalho de pesquisa de doutorado estão enumeradas a seguir (as contribuições consideradas mais relevantes estão ressaltadas em **negrito**):

1. **Um conjunto de aproximações podadas para a DTT foi derivado;**
2. As transformadas inversas para as aproximações podadas para a DTT foram analisadas.

As matrizes transpostas apresentaram menor sobrecarga computacional, assim como me-

nor desvio de diagonalidade quando comparadas com inversas generalizadas;

3. As aproximações podadas propostas para a DTT foram empregadas em sistemas de codificação tipo JPEG e codecs H.264/AVC e H.265/HEVC. Os resultados apresentaram desempenho competitivo a um custo computacional consideravelmente menor;
4. Arquiteturas em FPGA para os métodos podados propostos foram desenvolvidas. Os métodos derivados apresentaram uma considerável economia em termos de potência e recursos de *hardware*;
5. Uma formulação algébrica generalizada para aproximações multidimensionais para a DCT foi proposta;
6. **Um conjunto de aproximações para a 3D DCT de dimensão $8 \times 8 \times 8$ foi derivado;**
7. A complexidade computacional para o caso generalizado foi discutida e formulada algebricamente;
8. Uma análise de *trade-off* foi desenvolvida, em que uma função custo foi proposta em termos de métricas de desempenho e complexidade. Todos os cenários realísticos foram cobertos pela análise e as aproximações propostas foram os métodos ótimos para a maioria dos casos;
9. Um procedimento de modificação de um dado volume de quantização 3D foi introduzido visando explorar as características de baixa complexidade das aproximações propostas para a 3D DCT;
10. Uma simulação de compressão de vídeo empregando os métodos aproximados propostos para a 3D DCT foi realizada. Ambas as avaliações quantitativas e qualitativas sugerem que os métodos propostos apresentam *trade-off* favorável entre desempenho e complexidade;
11. A aproximação 3D MRDCT obtida foi incorporada a um sistema de rastreamento visual baseado na 3D DCT exata. Avaliações qualitativas e quantitativas mostraram desempenhos muito próximos.

12. Um conjunto de aproximações quase-ortogonais para a DHT de comprimento oito foi introduzido, considerando medidas de complexidade computacional, desempenho e quase-ortogonalidade;
13. Foram analisadas alternativas para o cômputo da transformada inversa de modo a reduzir o desvio de diagonalidade apresentado pelas aproximações propostas;
14. Foi proposta uma formulação algébrica para derivar aproximações para a 3D DHT que considera as propriedades de não-separabilidade da transformada exata.
15. **Diversos métodos aproximados para a 3D DHT foram propostos baseados nas matrizes aproximadas derivadas e no formalismo proposto e a complexidade computacional foi analisada;**
16. Os métodos aproximados derivados para a 3D DHT foram considerados em um experimento de compressão tridimensional de imagens médicas. Os resultados qualitativos e quantitativos sugerem desempenhos praticamente iguais a um custo computacional consideravelmente menor.

5.2 TRABALHOS FUTUROS

Como sugestão de trabalhos futuros baseados no presente trabalho, destacam-se as seguintes linhas de pesquisa: expandir a abordagem de transformadas podadas para o caso multidimensional, derivar transformadas aproximadas de comprimentos mais longos, aplicar transformadas aproximadas de dimensões mais elevadas em aplicações práticas, desenvolver arquiteturas de *hardware* para os métodos aproximados multidimensionais. As sugestões de trabalhos futuros são detalhadas a seguir:

- Buscar por transformadas aproximadas de comprimentos mais longos: A tarefa de derivar matrizes aproximadas se torna acentuadamente mais desafiadora quando comprimentos de transformada mais longos são considerados (ex. $N = 16, 32, 64, 128, 256$). Este fato está associado com o incremento substancial no número de parâmetros envolvidos, tendo como consequência um maior espaço de busca. Neste contexto, abordagens escalonadas

são uma alternativa, em que se explora uma fatoração da transformada de comprimento N em produtos que dependem de transformadas de comprimento menor.

- Derivar aproximações para a 4D DCT visando compressão de dados de campo de luz: Dados de campo de luz são arranjos quadrimencionais altamente correlacionados (KAMAL et al., 2016; CARVALHO et al., 2018; MAGNOR; GIROD, 2000; LEVOY; HANRAHAN, 1996). O volume de dados para armazenamento é notavelmente alto, assim como a quantidade de operações aritméticas necessárias para realizar compressão dos dados. Sendo assim, ferramentas de baixo custo computacional para compressão de dados são desejáveis. Aproximações de baixa complexidade para a 4D DCT podem ser derivadas aplicando o formalismo generalizado proposto neste trabalho empregando novas matrizes aproximadas otimizadas para este contexto.
- Estender a abordagem de transformadas podadas para o caso multidimensional: A propriedade de compactação da energia é herdada pelas transformadas multidimensionais e algoritmos podados podem ser considerados para uma redução acentuada no custo computacional de aplicações de compressão de imagem baseadas em transformadas multidimensionais;
- Desenvolver arquiteturas em FPGA para os métodos aproximados multidimensionais derivados, visando sistemas de compressão de vídeo com processamento em tempo real e baixo consumo energético.

REFERÊNCIAS

- ABDI, H.; WILLIAMS, L. J. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, Wiley Online Library, v. 2, n. 4, p. 433–459, 2010. Citado na página 89.
- ABOUSLEMAN, G. P.; MARCELLIN, M. W.; HUNT, B. R. Compression of hyperspectral imagery using the 3-D DCT and hybrid DPCM/DCT. *IEEE Transactions on Geoscience and Remote Sensing*, v. 33, n. 1, p. 26–34, 1995. Citado 2 vezes nas páginas 81 e 84.
- ABRAMOWITZ, M.; STEGUN, I. Handbook of mathematical functions: with formulas, graphs, and mathematical tables, 9th printing dover publication. *New York*, 1970. Citado 2 vezes nas páginas 31 e 38.
- ABU, N. A. et al. An efficient compact Tchebichef moment for image compression. In: *IEEE. 10th International Conference on Information Sciences Signal Processing and their Applications (ISSPA)*. [S.l.], 2010. p. 448–451. Citado 3 vezes nas páginas 24, 31 e 32.
- AGBINYA, J. I. Fast interpolation algorithm using fast Hartley transform. *Proceedings of the IEEE*, v. 75, n. 4, p. 523–524, 1987. Citado 2 vezes nas páginas 25 e 94.
- AHMED, N.; NATARAJAN, T.; RAO, K. R. Discrete cosine transform. *IEEE Transactions on Computers*, C-23, p. 90–93, 1974. ISSN 0018-9340. Citado 3 vezes nas páginas 20, 66 e 91.
- AHMED, N.; RAO, K. R. *Orthogonal Transforms for Digital Signal Processing*. [S.l.]: Springer, 1975. Citado 2 vezes nas páginas 19 e 94.
- AL-HILO, E. A.; GEORGE, L. E. Speeding-up fractal colored image compression using moments features. In: *IEEE. Digital Image Computing: Techniques and Applications (DICTA)*. [S.l.], 2008. p. 486–490. Citado na página 31.
- ALMAGAMBETOV, A.; VELIPASALAR, S.; CASARES, M. Robust and computationally lightweight autonomous tracking of vehicle taillights and signal detection by embedded smart cameras. *IEEE Transactions on Industrial Electronics*, v. 62, n. 6, p. 3732–3741, 2015. Citado 2 vezes nas páginas 89 e 90.
- ANITHA, S.; NAGABHUSHANA, B. S. Design and development of a novel algorithm for 4D image compression. *Digital Image Processing*, v. 3, n. 20, p. 1264–1268, 2011. Citado na página 27.
- ARAI, Y.; AGUI, T.; NAKAJIMA, M. A fast DCT-SQ scheme for images. *Transactions of the IEICE*, E-71, p. 1095–1097, 1988. Citado na página 21.
- BÄCKSTRÖM, T. Decorrelating MVDR filterbanks using the non-uniform discrete Fourier transform. *IEEE Signal Processing Letters*, IEEE, v. 22, n. 4, p. 479–483, 2015. Citado na página 19.
- BARJATYA, A. Block matching algorithms for motion estimation. *IEEE Transactions Evolution Computation*, v. 8, n. 3, p. 225–239, 2004. Citado na página 86.
- BARRÉ, S. *Medical Image Samples*. <http://www.barre.nom.fr/medical/samples/>. Disponível em: <<http://www.barre.nom.fr/medical/samples/>>. Citado na página 106.

- BAYER, F. M.; CINTRA, R. J. DCT-like transform for image compression requires 14 additions only. *Electronics Letters*, v. 48, p. 919–921, 2012. ISSN 0013-5194. Citado 7 vezes nas páginas 21, 22, 23, 66, 72, 73 e 74.
- BELKASIM, S. O.; SHRIDHAR, M.; AHMADI, M. Pattern recognition with moment invariants: a comparative study and new results. *Pattern recognition*, Elsevier, v. 24, n. 12, p. 1117–1138, 1991. Citado na página 31.
- BENESTY, J. et al. Pearson correlation coefficient. In: *Noise reduction in speech processing*. [S.l.]: Springer, 2009. p. 1–4. Citado na página 88.
- BERNSTEIN, D. S. *Matrix Mathematics: Theory, Facts, and Formulas*. [S.l.]: Princeton University Press, 2009. Citado 9 vezes nas páginas 35, 36, 39, 40, 42, 43, 44, 68 e 97.
- BHARATI, M. H.; LIU, J. J.; MACGREGOR, J. F. Image texture analysis: methods and comparisons. *Chemometrics and intelligent laboratory systems*, Elsevier, v. 72, n. 1, p. 57–71, 2004. Citado na página 88.
- BHASKARAN, V.; KONSTANTINIDES, K. *Image and Video Compression Standards*. [S.l.]: Kluwer Academic Publishers, 1997. Citado 8 vezes nas páginas 19, 20, 51, 52, 81, 82, 85 e 107.
- BLAHUT, R. E. *Fast Algorithms for Signal Processing*. [S.l.]: Cambridge University Press, 2010. Citado 6 vezes nas páginas 19, 20, 78, 79, 91 e 106.
- BORS, A. G. Watermarking mesh-based representations of 3-d objects using local moments. *IEEE Transactions on Image processing*, v. 15, n. 3, p. 687–701, 2006. Citado na página 31.
- BOSSSEN, F. et al. HEVC complexity and implementation analysis. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 22, n. 12, p. 1685–1696, 2012. Citado na página 41.
- BOUGUEZEL, S.; AHMAD, M. O.; SWAMY, M. N. S. Efficient pruning algorithms for the DFT computation for a subset of output samples. In: IEEE. *Circuits and Systems, 2003. ISCAS'03. Proceedings of the 2003 International Symposium on*. [S.l.], 2003. v. 4, p. IV–IV. Citado na página 23.
- BOUGUEZEL, S.; AHMAD, M. O.; SWAMY, M. N. S. Low-complexity 8×8 transform for image compression. *Electronics Letters*, v. 44, p. 1249–1250, 2008. ISSN 0013-5194. Citado 9 vezes nas páginas 21, 22, 23, 41, 66, 72, 73, 74 e 107.
- BOUGUEZEL, S.; AHMAD, M. O.; SWAMY, M. N. S. A fast 8×8 transform for image compression. In: *International Conference on Microelectronics (ICM)*. [S.l.: s.n.], 2009. p. 74–77. Citado 7 vezes nas páginas 21, 22, 23, 66, 72, 73 e 74.
- BOUGUEZEL, S.; AHMAD, M. O.; SWAMY, M. N. S. A novel transform for image compression. In: *53rd IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*. [S.l.: s.n.], 2010. p. 509–512. ISSN 1548-3746. Citado na página 23.
- BOUGUEZEL, S.; AHMAD, M. O.; SWAMY, M. N. S. Binary discrete cosine and Hartley transforms. *IEEE Transactions on Circuits and Systems I: Regular Papers*, v. 60, p. 989–1002, 2013. Citado 9 vezes nas páginas 21, 22, 23, 27, 66, 72, 73, 74 e 95.

BOUSSAKTA, S. Fast algorithm for the 3-D DCT-II. *IEEE Transactions on Signal Processing*, v. 52, p. 992–1001, 2004. Citado 6 vezes nas páginas 20, 26, 67, 68, 76 e 77.

BOUSSAKTA, S.; ALSHIBAMI, O. H.; AZIZ, M. Y. Radix- $2 \times 2 \times 2$ algorithm for the 3-D discrete Hartley transform. *IEEE transactions on signal processing*, v. 49, n. 12, p. 3145–3156, 2001. Citado na página 106.

BOZINOVIĆ, N.; KONRAD, J. Scan order and quantization for 3D-DCT coding. In: *Proceedings of SPIE Visual Communications and Image Processing*. [S.l.: s.n.], 2003. v. 5150, p. 1205. Citado 4 vezes nas páginas 26, 28, 82 e 114.

BOZINOVIĆ, N.; KONRAD, J. Motion analysis in 3D DCT domain and its application to video coding. *Signal Processing: Image Communication*, v. 20, p. 510–528, 2005. Citado 4 vezes nas páginas 20, 26, 28 e 82.

BRACEWELL, R. et al. Optical synthesis of the Hartley transform. *Applied optics*, Optical Society of America, v. 24, n. 10, p. 1401–1402, 1985. Citado na página 94.

BRACEWELL, R. N. Discrete Hartley transform. *Journal of the Optical Society of America*, Optical Society of America, v. 73, n. 12, p. 1832–1835, 1983. Citado 3 vezes nas páginas 25, 94 e 96.

BRACEWELL, R. N. The fast Hartley transform. *Proceedings of the IEEE*, v. 72, n. 8, p. 1010–1018, 1984. Citado 3 vezes nas páginas 25, 94 e 95.

BRIGGS, W. L.; HENSON, V. E. *The DFT: an owners' manual for the discrete Fourier transform*. [S.l.]: Siam, 1995. Citado 2 vezes nas páginas 78 e 79.

BRITANAK, V.; YIP, P.; RAO, K. R. *Discrete Cosine and Sine Transforms*. [S.l.]: Academic Press, 2007. Citado 12 vezes nas páginas 19, 20, 21, 22, 39, 40, 41, 73, 78, 82, 90 e 105.

BUDAGAVI, M. et al. Core transform design in the high efficiency video coding (HEVC) standard. *IEEE Journal of Selected Topics in Signal Processing*, v. 7, n. 6, p. 1029–1041, 2013. ISSN 1932-4553. Citado na página 58.

CAO, S. et al. A Hadamard product based method for DOA estimation and gain-phase error calibration. *IEEE Transactions on Aerospace and Electronic Systems*, v. 49, p. 1224–1233, 2013. Citado na página 40.

CARDOSO, G.; SANIIE, J. Performance evaluation of DWT, DCT, and WHT for compression of ultrasonic signals. In: *IEEE Ultrasonics Symposium*. [S.l.: s.n.], 2004. v. 3, p. 2314–2317. Citado na página 26.

CARVALHO, M. B. de et al. A 4D DCT-based lenslet light field codec. In: *IEEE. 25th International Conference on Image Processing (ICIP)*. [S.l.], 2018. p. 435–439. Citado 2 vezes nas páginas 27 e 117.

CHAN, R. K. W.; LEE, M. C. 3D-DCT quantization as a compression technique for video sequences. *Proceedings of the International Conference on Virtual Systems and MultiMedia*, p. 188–196, 1997. Citado 6 vezes nas páginas 25, 26, 28, 65, 66 e 82.

CHAN, Y.-L.; SIU, W.-C. Variable temporal-length 3-D discrete cosine transform coding. *IEEE Transactions on Image Processing*, IEEE, v. 6, n. 5, p. 758–763, 1997. Citado 4 vezes nas páginas 26, 28, 66 e 114.

CHANG, C.-Y.; LIE, H. W. Real-time visual tracking and measurement to control fast dynamics of overhead cranes. *IEEE Transactions on Industrial Electronics*, v. 59, n. 3, p. 1640–1649, 2012. Citado na página 90.

CHEN, L.; ZHAO, D. Optical image encryption with Hartley transforms. *Optics letters*, Optical Society of America, v. 31, n. 23, p. 3438–3440, 2006. Citado na página 94.

CHEN, W. H.; SMITH, C.; FRALICK, S. A fast computational algorithm for the discrete cosine transform. *IEEE Transactions on Communications*, v. 25, p. 1004–1009, 1977. ISSN 0090-6778. Citado 2 vezes nas páginas 21 e 77.

CHEN, X.; DAI, Q.; LI, C. A fast algorithm for computing multidimensional DCT on certain small sizes. *IEEE Transactions on Signal Processing*, IEEE, v. 51, n. 1, p. 213–220, 2003. Citado na página 27.

CHEN, Y.-J.; ORAINTARA, S.; NGUYEN, T. Video compression using integer DCT. In: *IEEE Proceedings of International Conference on Image Processing*. [S.l.: s.n.], 2000. v. 2, p. 844–845. Citado na página 21.

CHEN, Y.-J. et al. Multiplierless approximation of transforms using lifting scheme and coordinate descent with adder constraint. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. [S.l.: s.n.], 2002. v. 3, p. III–3136. Citado na página 21.

CHENG, P.-W.; SHEN, C.-C.; LI, P.-C. MPEG compression of ultrasound RF channel data for a real-time software-based imaging system. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, IEEE, v. 59, n. 7, p. 1413–1420, 2012. Citado na página 26.

CHO, N. Fast algorithm and implementation of 2-D discrete cosine transform. *IEEE Transactions on Circuits and Systems*, v. 38, p. 297–305, 1991. Citado 2 vezes nas páginas 20 e 67.

CHONG, C.-W.; RAVEENDRAN, P.; MUKUNDAN, R. Translation and scale invariants of Legendre moments. *Pattern recognition*, Elsevier, v. 37, n. 1, p. 119–129, 2004. Citado na página 31.

CINTRA, R. J. An integer approximation method for discrete sinusoidal transforms. *Journal of Circuits, Systems, and Signal Processing*, v. 30, p. 1481–1501, 2011. Citado 7 vezes nas páginas 21, 22, 27, 41, 66, 74 e 95.

CINTRA, R. J.; BAYER, F. M. A DCT approximation for image compression. *IEEE Signal Processing Letters*, v. 18, p. 579–582, 2011. Citado 8 vezes nas páginas 21, 22, 23, 66, 72, 73, 74 e 90.

CINTRA, R. J. et al. Energy-efficient 8-point DCT approximations: Theory and hardware architectures. *Circuits, Systems, and Signal Processing*, Springer, v. 35, n. 11, p. 4009–4029, 2016. Citado 7 vezes nas páginas 22, 23, 43, 44, 45, 46 e 78.

CINTRA, R. J.; BAYER, F. M.; TABLADA, C. J. Low-complexity 8-point DCT approximations based on integer functions. *Signal Processing*, v. 99, p. 201–214, 2014. ISSN 0165-1684. Citado 12 vezes nas páginas 21, 22, 39, 40, 41, 42, 43, 66, 72, 73, 74 e 99.

- COIFMAN, B. et al. A real-time computer vision system for vehicle tracking and traffic surveillance. *Transportation Research Part C: Emerging Technologies*, Elsevier, v. 6, n. 4, p. 271–288, 1998. Citado na página 89.
- COMPUTER Vision Lab. SNU Database. Disponível em: <<http://cv.snu.ac.kr/research/>>. Citado na página 92.
- COUTINHO, V. A. *Aproximações para DCT via pruning com aplicações em codificação de imagem e vídeo*. [S.l.]: Dissertação de Mestrado. Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Pernambuco, 2015. Citado na página 58.
- COUTINHO, V. A.; CINTRA, R. J.; BAYER, F. M. Low-complexity multidimensional DCT approximations for high-order tensor data decorrelation. *IEEE Transactions on Image Processing*, v. 26, n. 5, p. 2296–2310, 2017. Citado 10 vezes nas páginas 76, 80, 81, 86, 87, 88, 89, 90, 91 e 105.
- COUTINHO, V. A. et al. Low-complexity pruned 8-point DCT approximations for image encoding. In: *IEEE International Conference on Electronics, Communications and Computers (CONIELECOMP)*. [S.l.: s.n.], 2015. p. 1–7. Citado 2 vezes nas páginas 23 e 44.
- COUTINHO, V. A. et al. A multiplierless pruned DCT-like transformation for image and video compression that requires ten additions only. *Journal of Real-Time Image Processing*, Springer, v. 12, n. 2, p. 247–255, 2016. Citado 3 vezes nas páginas 22, 23 e 25.
- COUTINHO, V. A. et al. Pruned discrete Tchebichef transform approximation for image compression. *Circuits, Systems, and Signal Processing*, Springer, p. 1–21, 2018. Citado 12 vezes nas páginas 45, 47, 49, 50, 53, 54, 55, 56, 57, 62, 63 e 64.
- DAI, Q.; CHEN, X.; LIN, C. Fast algorithms for multidimensional DCT-to-DCT computation between a block and its associated subblocks. *IEEE Transactions on Signal Processing*, IEEE, v. 53, n. 8, p. 3219–3225, 2005. Citado na página 27.
- DANCIU, M. et al. 3D DCT supervised segmentation applied on liver volumes. In: *IEEE. 35th International Conference on Telecommunications and Signal Processing (TSP)*. [S.l.], 2012. p. 779–783. Citado 2 vezes nas páginas 20 e 26.
- DAWOOD, A. M.; GHANBARI, M. Content-based MPEG video traffic modeling. *IEEE Transactions on Multimedia*, v. 1, n. 1, p. 77–87, 1999. Citado 2 vezes nas páginas 86 e 88.
- DEMIGHA, O.; HIDOUCI, W.-K.; AHMED, T. On energy efficiency in collaborative target tracking in wireless sensor network: a review. *IEEE Communications Surveys & Tutorials*, v. 15, n. 3, p. 1210–1222, 2013. Citado na página 89.
- DESCHAMPS, J.-P.; BIOUL, G. J.; SUTTER, G. D. *Synthesis of arithmetic circuits: FPGA, ASIC and embedded systems*. [S.l.]: John Wiley & Sons, 2006. Citado na página 19.
- DESCOUR, M. R. et al. Toward the development of miniaturized imaging systems for detection of pre-cancer. *IEEE Journal of Quantum Electronics*, v. 38, n. 2, p. 122–130, 2002. Citado na página 95.
- DESMOULIERS, C.; ORUKLU, E.; SANIIE, J. Adaptive 3D ultrasonic data compression using distributed processing engines. In: *IEEE International Ultrasonics Symposium (IUS)*. [S.l.: s.n.], 2009. p. 694–697. Citado na página 26.

- DING, H.; SOON, Y.; YEO, C. K. A DCT-based speech enhancement system with pitch synchronous analysis. *IEEE Transactions on audio, speech, and language processing*, IEEE, v. 19, n. 8, p. 2614–2623, 2011. Citado na página 19.
- DONG, Y. et al. Multiscale sampling based texture image classification. *IEEE Signal Processing Letters*, IEEE, v. 24, n. 5, p. 614–618, 2017. Citado na página 88.
- DUDGEON, D. E.; MERSEREAU, R. M. *Multidimensional Digital Signal Processing*. [S.l.]: Prentice Hall Professional Technical Reference, 1990. Citado na página 32.
- DUFAUX, F.; KONRAD, J. Efficient, robust, and fast global motion estimation for video coding. *IEEE transactions on image processing*, v. 9, n. 3, p. 497–501, 2000. Citado na página 65.
- DUFAUX, F.; MOSCHENI, F. Motion estimation techniques for digital TV: A review and a new contribution. *Proceedings of the IEEE*, v. 83, n. 6, p. 858–876, 1995. Citado 2 vezes nas páginas 65 e 66.
- DULEBA, I. Hartley transform in compression of medical ultrasonic images. In: IEEE. *Proceedings of the International Conference on Image Analysis and Processing*. [S.l.], 1999. p. 722–727. Citado 2 vezes nas páginas 27 e 95.
- FAN, C.-P.; SU, G.-A. Pruning fast Fourier transform algorithm design using group-based method. *Signal Processing*, Elsevier, v. 87, n. 11, p. 2781–2798, 2007. Citado na página 23.
- FAN, Y. et al. A hardware-oriented IME algorithm for HEVC and its hardware implementation. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 28, n. 8, p. 2048–2057, 2018. Citado na página 65.
- FEIG, E.; WINOGRAD, S. Fast algorithms for the discrete cosine transform. *IEEE Transactions on Signal Processing*, v. 40, p. 2174–2193, 1992. Citado na página 21.
- FITZEK, F. H.; REISSLEIN, M. MPEG-4 and H.263 video traces for network performance evaluation. *IEEE Network*, v. 15, n. 6, p. 40–54, 2001. Citado na página 84.
- FLUSSER, J.; SUK, T. Pattern recognition by affine moment invariants. *Pattern recognition*, v. 26, n. 1, p. 167–174, 1993. Citado na página 31.
- FONG, C.-K.; CHAM, W.-K. LLM integer cosine transform and its fast algorithm. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 22, p. 844–854, 2012. Citado na página 21.
- GALL, D. L. MPEG: A video compression standard for multimedia applications. *Communications of the ACM*, ACM, v. 34, n. 4, p. 46–58, 1991. Citado na página 25.
- GAO, Y. et al. Low-power ultrawideband wireless telemetry transceiver for medical sensor applications. *IEEE Transactions on Biomedical Engineering*, IEEE, v. 58, n. 3, p. 768–772, 2011. Citado na página 95.
- GERBRANDS, J. J. On the relationships between SVD, KLT and PCA. *Pattern Recognition*, Elsevier, v. 14, n. 1, p. 375–381, 1981. Citado na página 89.
- GEREK, O. N.; ÇETIN, A. E. A 2-D orientation-adaptive prediction filter in lifting structures for image coding. *IEEE Transactions on Image Processing*, v. 15, n. 1, p. 106–111, 2006. Citado na página 41.

- GESTER, D.; SIMON, S. A spatial moments sub-pixel edge detector with edge blur compensation for imaging metrology. In: IEEE. *International Instrumentation and Measurement Technology Conference (I2MTC)*. [S.l.], 2018. Citado na página 31.
- GHARAVI, H.; MILLS, M. Blockmatching motion estimation algorithms-new results. *IEEE Transactions on Circuits and Systems*, v. 37, n. 5, p. 649–651, 1990. Citado na página 65.
- GHOSAL, S.; MEHROTRA, R. Orthogonal moment operators for subpixel edge detection. *Pattern recognition*, Elsevier, v. 26, n. 2, p. 295–306, 1993. Citado na página 31.
- GONZALEZ, R. C.; WOODS, R. *Digital Image Processing*. [S.l.]: Prentice Hall, 2001. Citado 7 vezes nas páginas 20, 25, 27, 34, 35, 94 e 95.
- GORDON, S.; MARPE, D.; WIEGAND, T. *Simplified use of 8x8 transform—Updated proposal and results*. 2004. Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, doc. JVT-K028, Munich, Germany. Citado na página 56.
- GORJI, H.; HADDADNIA, J. A novel method for early diagnosis of Alzheimer’s disease based on pseudo Zernike moment from structural MRI. *Neuroscience*, Elsevier, v. 305, p. 361–371, 2015. Citado na página 31.
- GORTLER, S. J. et al. The lumigraph. In: ACM. *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. [S.l.], 1996. p. 43–54. Citado na página 27.
- GOVINDAN, P. et al. Hardware and software architectures for computationally efficient three-dimensional ultrasonic data compression. *IET Circuits, Devices & Systems*, IET, v. 10, n. 1, p. 54–61, 2016. Citado na página 26.
- GRIGORYAN, A. M. A novel algorithm for computing the 1-D discrete Hartley transform. *IEEE Signal Processing Letters*, v. 11, n. 2, p. 156–159, 2004. Citado 2 vezes nas páginas 25 e 95.
- GRINSHPAN, A. Z. Generalized hypergeometric functions: product identities and weighted norm inequalities. *The Ramanujan Journal*, Springer, v. 31, n. 1-2, p. 53–66, 2013. Citado na página 34.
- GUO, J.-I. An efficient design for one-dimensional discrete Hartley transform using parallel additions. *IEEE transactions on signal processing*, IEEE, v. 48, n. 10, p. 2806–2813, 2000. Citado 2 vezes nas páginas 25 e 95.
- HADDADNIA, J.; FAEZ, K.; MOALLEM, P. Neural network based face recognition with moment invariants. In: IEEE. *Image Processing, 2001. Proceedings. 2001 International Conference on*. [S.l.], 2001. v. 1, p. 1018–1021. Citado na página 31.
- HAO, H.; BRACEWELL, R. N. A three-dimensional DFT algorithm using the fast Hartley transform. *Proceedings of the IEEE*, v. 75, n. 2, p. 264–266, 1987. Citado na página 103.
- HARMUTH, H. F. *Transmission of information by orthogonal functions*. [S.l.]: Springer Science & Business Media, 2012. Citado na página 19.
- HARTLEY, R. Optimization of canonic signed digit multipliers for filter design. In: IEEE. *Circuits and Systems, 1991., IEEE International Symposium on*. [S.l.], 1991. p. 1992–1995. Citado na página 24.

- HARTLEY, R. V. A more symmetrical Fourier analysis applied to transmission problems. *Proceedings of the IRE, IEEE*, v. 30, n. 3, p. 144–150, 1942. Citado na página 94.
- HAWHEEL, T. I. A new square wave transform based on the DCT. *Signal Processing*, v. 82, p. 2309–2319, 2001. Citado 9 vezes nas páginas 21, 22, 23, 41, 42, 66, 72, 73 e 74.
- HEIDEMAN, M. T. *Multiplicative Complexity, Convolution, and the DFT*. [S.l.]: Springer-Verlag, 1988. (Signal Processing and Digital Filtering). ISBN 978-0-387-96810-0. Citado 5 vezes nas páginas 19, 21, 22, 77 e 94.
- HEYWOOD, M.; NOAKES, P. Fractional central moment method for movement-invariant object classification. *IEE Proceedings-Vision, Image and Signal Processing, IET*, v. 142, n. 4, p. 213–219, 1995. Citado na página 31.
- HNATIV, L. Integer cosine transforms: Methods to construct new order 8, 16 fast transforms and their application. *Cybernetics and Systems Analysis, Springer*, v. 50, n. 6, p. 913–929, 2014. Citado na página 21.
- HORN, R. A.; JOHNSON, C. R. *Matrix analysis*. [S.l.]: Cambridge university press, 2012. Citado na página 72.
- HOU, H. S. The fast Hartley transform algorithm. *IEEE Transactions on Computers, IEEE*, v. 100, n. 2, p. 147–156, 1987. Citado 2 vezes nas páginas 25 e 95.
- HOU, H. S. A fast recursive algorithm for computing the discrete cosine transform. *IEEE Transactions on Acoustic, Signal, and Speech Processing*, v. 35, p. 1455–1461, 1987. Citado na página 21.
- HSU, H.-S.; TSAI, W.-H. Moment-preserving edge detection and its application to image data compression. *Optical engineering, International Society for Optics and Photonics*, v. 32, n. 7, p. 1596–1609, 1993. Citado na página 31.
- HU, Y.; SAWAN, M. A fully integrated low-power BPSK demodulator for implantable medical devices. *IEEE Transactions on Circuits and Systems I: Regular Papers*, v. 52, n. 12, p. 2552–2562, 2005. Citado na página 95.
- HUANG, H. et al. Blind integrity verification of medical images. *IEEE transactions on information technology in biomedicine*, v. 16, n. 6, p. 1122–1126, 2012. Citado na página 32.
- HUNT, O.; MUKUNDAN, R. A comparison of discrete orthogonal basis functions for image compression. University of Canterbury. Computer Science and Software Engineering., 2004. Citado 3 vezes nas páginas 24, 32 e 33.
- HUSSMANN, S.; LIEPERT, T. Three-dimensional TOF robot vision system. *IEEE Transactions on Instrumentation and Measurement*, v. 58, n. 1, p. 141–146, 2009. Citado na página 31.
- International Organisation for Standardisation. *Generic Coding of Moving Pictures and Associated Audio Information – Part 2: Video*. [S.l.], 1994. Citado 3 vezes nas páginas 21, 77 e 81.
- International Telecommunication Union. *ITU-T Recommendation H.261 Version 1: Video Codec for Audiovisual Services at $p \times 64$ kbits*. [S.l.], 1990. Citado 2 vezes nas páginas 21 e 81.

International Telecommunication Union. *ITU-T Recommendation H.263 version 1: Video Coding for Low Bit Rate Communication*. [S.l.], 1995. Citado 2 vezes nas páginas 21 e 81.

International Telecommunication Union. *High efficiency video coding: Recommendation ITU-T H.265*. [S.l.], 2013. Citado 5 vezes nas páginas 21, 22, 49, 77 e 82.

ISHWAR, S.; MEHER, P. K.; SWAMY, M. N. S. Discrete Tchebichef transform-a fast 4×4 algorithm and its application in image/video compression. In: IEEE. *International Symposium on Circuits and Systems (ISCAS)*. [S.l.], 2008. p. 260–263. Citado 3 vezes nas páginas 24, 32 e 38.

JACOB, J. A.; KUMAR, N. S. FPGA implementation of optimal 3D-integer DCT structure for video compression. *The Scientific World Journal*, Hindawi Publishing Corporation, v. 2015, 2015. Citado 2 vezes nas páginas 26 e 27.

JAKUBCZAK, S.; KATABI, D. Softcast: Clean-slate scalable wireless video. In: *48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. [S.l.: s.n.], 2010. p. 530–533. Citado 2 vezes nas páginas 26 e 114.

JAKUBCZAK, S.; KATABI, D. Softcast: one-size-fits-all wireless video. *ACM SIGCOMM Computer Communication Review*, ACM, v. 41, n. 4, p. 449–450, 2011. Citado 2 vezes nas páginas 26 e 114.

JAKUBCZAK, S.; RAHUL, H.; KATABI, D. SoftCast: one-size-fits-all wireless video. *ACM SIGCOMM Computer Communication Review*, v. 40, p. 449–450, 2010. Citado 2 vezes nas páginas 26 e 114.

JAVED, O.; SHAH, M. Tracking and object classification for automated surveillance. In: *Computer Vision–ECCV 2002*. [S.l.]: Springer, 2002. p. 343–357. Citado na página 89.

Jl, Z. et al. A moment-based nonlocal-means algorithm for image denoising. *Information Processing Letters*, Elsevier, v. 109, n. 23-24, p. 1238–1244, 2009. Citado na página 31.

JUEFEI-XU, F.; SAVVIDES, M. Subspace-based discrete transform encoded local binary patterns representations for robust periocular matching on nist's face recognition grand challenge. *IEEE transactions on image processing*, IEEE, v. 23, n. 8, p. 3490–3505, 2014. Citado na página 19.

KAFAI, M.; ESHGHI, K.; BHANU, B. Discrete cosine transform locality-sensitive hashes for face retrieval. *IEEE Transactions on Multimedia*, IEEE, v. 16, n. 4, p. 1090–1103, 2014. Citado na página 19.

KAMAL, M. H. et al. Tensor low-rank and sparse light field photography. *Computer Vision and Image Understanding*, Elsevier, v. 145, p. 172–181, 2016. Citado 2 vezes nas páginas 27 e 117.

KHOTANZAD, A.; HONG, Y. H. Invariant image recognition by Zernike moments. *IEEE Transactions on pattern analysis and machine intelligence*, v. 12, n. 5, p. 489–497, 1990. Citado na página 31.

KIM, H. S.; LEE, H.-K. Invariant image watermark using Zernike moments. *IEEE transactions on Circuits and Systems for Video Technology*, v. 13, n. 8, p. 766–775, 2003. Citado na página 31.

- KIRBY, M.; SIROVICH, L. Application of the Karhunen-Loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 12, p. 103–108, 1990. Citado na página 19.
- KOK, C. W. Fast algorithm for computing discrete cosine transform. *IEEE Transactions on Signal Processing*, v. 45, p. 757 – 760, 1997. Citado na página 76.
- KOUADRIA, N. et al. Low complexity DCT for image compression in wireless visual sensor networks. *Electronics Letters*, v. 49, p. 1531–1532, 2013. Citado 2 vezes nas páginas 22 e 25.
- KOUADRIA, N. et al. Pruned discrete Tchebichef transform for image coding in wireless multimedia sensor networks. *AEU-International Journal of Electronics and Communications*, Elsevier, v. 74, p. 123–127, 2017. Citado 3 vezes nas páginas 24, 33 e 43.
- LAI, T. H.; GUAN, L. Video coding algorithm using 3-D DCT and vector quantization. In: *Proceedings of the International Conference on Image Processing*. [S.l.: s.n.], 2002. v. 1, p. I-741. Citado 6 vezes nas páginas 26, 28, 80, 81, 84 e 114.
- LATHAUWER, L. D.; MOOR, B. D. From matrix to tensor: Multilinear algebra and signal processing. In: CITESEER. *Institute of Mathematics and Its Applications Conference Series*. [S.l.], 1998. v. 67, p. 1–16. Citado 2 vezes nas páginas 28 e 68.
- LATHAUWER, L. D.; MOOR, B. D.; VANDEWALLE, J. On the best rank-1 and rank-(r_1, r_2, \dots, r_n) approximation of higher-order tensors. *SIAM Journal on Matrix Analysis and Applications*, SIAM, v. 21, n. 4, p. 1324–1342, 2000. Citado 4 vezes nas páginas 28, 68, 69 e 70.
- LECUIRE, V.; MAKKAOUI, L.; MOUREAUX, J.-M. Fast zonal DCT for energy conservation in wireless image sensor networks. *Electronics Letters*, v. 48, p. 125–127, 2012. Citado 8 vezes nas páginas 22, 23, 25, 43, 52, 61, 77 e 78.
- LEE, B. G. A new algorithm for computing the discrete cosine transform. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. 32, p. 1243–1245, 1984. Citado na página 21.
- LEE, M.; CHAN, R. K.; ADJEROH, D. A. Quantization of 3D-DCT coefficients and scan order for video compression. *Journal of Visual Communication and Image Representation*, Elsevier, v. 8, n. 4, p. 405–422, 1997. Citado 5 vezes nas páginas 26, 28, 82, 85 e 114.
- LEE, M.-H. CSD filter design for VLSI implementation of GA-VSB receiver. *IEEE Transactions on Consumer Electronics*, v. 43, p. 197–206, 1997. Citado 2 vezes nas páginas 22 e 41.
- LENGWEHASATIT, K.; ORTEGA, A. Scalable variable complexity approximate forward DCT. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 14, p. 1236–1248, 2004. ISSN 1051-8215. Citado 10 vezes nas páginas 21, 22, 23, 66, 72, 73, 74, 76, 82 e 90.
- LEVOY, M.; HANRAHAN, P. Light field rendering. In: ACM. *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. [S.l.], 1996. p. 31–42. Citado 2 vezes nas páginas 27 e 117.
- LI, E. et al. Image projective invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, 2018. Citado na página 31.

- LI, L.; HOU, Z. Multiview video compression with 3D-DCT. In: *ITI 5th International Conference on Information and Communications Technology*. [S.l.: s.n.], 2007. p. 59–61. Citado 3 vezes nas páginas 80, 81 e 84.
- LI, L. et al. Referenceless measure of blocking artifacts by Tchebichef kernel analysis. *IEEE Signal Processing Letters*, IEEE, v. 21, n. 1, p. 122–125, 2014. Citado na página 32.
- LI, X. et al. Incremental learning of 3D-DCT compact representations for robust visual tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 35, p. 863–881, 2013. Citado 9 vezes nas páginas 26, 28, 67, 68, 71, 90, 91, 92 e 114.
- LI, X.; GUO, Q.; LU, X. Spatiotemporal statistics for video quality assessment. *IEEE Transactions on Image Processing*, IEEE, v. 25, n. 7, p. 3329–3342, 2016. Citado na página 26.
- LIANG, J.; TRAN, T. D. Fast multiplierless approximation of the DCT with the lifting scheme. *IEEE Transactions on Signal Processing*, v. 49, p. 3032–3044, 2001. Citado 4 vezes nas páginas 21, 82, 90 e 106.
- LIAO, S. X.; PAWLAK, M. On image analysis by moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 18, n. 3, p. 254–266, 1996. Citado na página 33.
- LIM, C.-L. et al. Plain, edge, texture (PET) block classifier using Tchebichef moments and SVM. In: IEEE. *International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS)*. [S.l.], 2013. p. 409–412. Citado na página 31.
- LIM, J. S. *Two-Dimensional Signal and Image Processing*. [S.l.]: Prentice Hall, 1990. Citado 2 vezes nas páginas 20 e 23.
- LIU, Z. et al. Color image encryption by using the rotation of color vector in Hartley transform domains. *Optics and Lasers in Engineering*, Elsevier, v. 48, n. 7, p. 800–805, 2010. Citado 2 vezes nas páginas 25 e 94.
- LOEFFLER, C.; LIGTENBERG, A.; MOSCHYTZ, G. S. Practical fast 1-D DCT algorithms with 11 multiplications. *ICASSP International Conference on Acoustics, Speech, and Signal Processing*, v. 2, p. 988–991, 1989. Citado 3 vezes nas páginas 21, 61 e 77.
- MADANAYAKE, A. et al. Low-power VLSI architectures for DCT/DWT: precision vs approximation for HD video, biomedical, and smart antenna applications. *IEEE Circuits and Systems Magazine*, v. 15, n. 1, p. 25–47, 2015. Citado 3 vezes nas páginas 21, 22 e 66.
- MADISETTI, V. K. *The Digital Signal Processing Handbook*. [S.l.]: CRC Press, 2009. Citado 2 vezes nas páginas 21 e 36.
- MAGNOR, M.; GIROD, B. Data compression for light-field rendering. *IEEE Transactions on Circuits and Systems for Video Technology*, IEEE, v. 10, n. 3, p. 338–343, 2000. Citado 2 vezes nas páginas 27 e 117.
- MAHARANA, G.; MEHER, P. Algorithm for efficient interpolation of real-valued signals using discrete Hartley transform. *Computers & electrical engineering*, Elsevier, v. 23, n. 2, p. 129–134, 1997. Citado 2 vezes nas páginas 25 e 94.

MAKKAOUI, L.; LECUIRE, V.; MOUREAUX, J. Fast zonal DCT-based image compression for wireless camera sensor networks. In: IEEE. *2nd International Conference on Image Processing Theory Tools and Applications (IPTA)*. [S.l.], 2010. p. 126–129. Citado 4 vezes nas páginas 22, 23, 52 e 61.

MALVAR, H. Fast computation of discrete cosine transform through fast Hartley transform. *Electronics Letters*, IET, v. 22, n. 7, p. 352–353, 1986. Citado na página 94.

MANDAL, J. K.; GHOSAL, S. K. Separable discrete Hartley transform based invisible watermarking for color image authentication (SDHTIWCIA). In: *Advances in Computing and Information Technology*. [S.l.]: Springer, 2013. p. 767–776. Citado 2 vezes nas páginas 25 e 95.

MANDAL, M. K.; ABOULNASR, T.; PANCHANATHAN, S. Image indexing using moments and wavelets. *IEEE Transactions on Consumer Electronics*, v. 42, n. 3, p. 557–565, 1996. Citado na página 31.

MARKANDEY, V. et al. Robot sensing techniques based on high-dimensional moment invariants and tensors. *IEEE transactions on robotics and automation*, v. 8, n. 2, p. 186–195, 1992. Citado na página 31.

MARKEL, J. FFT pruning. *IEEE Transactions on Audio and Electroacoustics*, v. 19, p. 305–311, 1971. Citado na página 23.

MECHOU EK, K. et al. Low complexity dct approximation for image compression in wireless image sensor networks. *Journal of Circuits, Systems and Computers*, World Scientific, v. 25, n. 08, p. 1650088, 2016. Citado na página 22.

MEDINA-MELENDEZ, M.; ARIAS-ESTRADA, M.; CASTRO, A. Input and/or output pruning of composite length FFTs using a DIF-DIT transform decomposition. *IEEE Transactions on Signal Processing*, IEEE, v. 57, n. 10, p. 4124–4128, 2009. Citado na página 23.

MEHER, P. K.; SRIKANTHAN, T.; PATRA, J. C. Scalable and modular memory-based systolic architectures for discrete Hartley transform. *IEEE Transactions on Circuits and Systems I: Regular Papers*, v. 53, n. 5, p. 1065–1077, 2006. Citado 3 vezes nas páginas 25, 94 e 95.

MOREOLO, M. S.; MUÑOZ, R.; JUNYENT, G. Novel power efficient optical OFDM based on Hartley transform for intensity-modulated direct-detection systems. *Journal of lightwave Technology*, IEEE, v. 28, n. 5, p. 798–805, 2010. Citado na página 94.

MUKUNDAN, R. Improving image reconstruction accuracy using discrete orthonormal moments. *UC Research Repository*, University of Canterbury. Computer Science and Software Engineering., 2003. Citado 3 vezes nas páginas 31, 33 e 38.

MUKUNDAN, R.; ONG, S.; LEE, P. A. Image analysis by Tchebichef moments. *IEEE Transactions on image Processing*, v. 10, n. 9, p. 1357–1364, 2001. Citado 5 vezes nas páginas 24, 31, 32, 33 e 38.

MULLA, A. et al. Image compression scheme based on zig-zag 3D-DCT and LDPC coding. In: *Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on)*. [S.l.: s.n.], 2014. p. 2380–2384. Citado 3 vezes nas páginas 26, 28 e 114.

MULTICOREWARE. *x265 HEVC Encoder / H.265 Video Codec*. 2017. <x265.org>. Citado na página 57.

NAKAGAKI, K.; MUKUNDAN, R. A fast 4×4 forward discrete Tchebichef transform algorithm. *IEEE Signal Processing Letters*, v. 14, n. 10, p. 684–687, 2007. Citado na página 38.

NATARAJAN, T. R.; AHMED, N. On interframe transform coding. *IEEE Transactions on Communications*, v. 25, n. 11, p. 1323–1329, 1977. Citado 7 vezes nas páginas 25, 28, 65, 66, 80, 81 e 84.

NIE, Y.; MA, K.-K. Adaptive rood pattern search for fast block-matching motion estimation. *IEEE Transactions on Image Processing*, v. 11, n. 12, p. 1442–1449, 2002. Citado na página 86.

NORTHCOTT, D. G. *Multilinear algebra*. [S.l.]: Cambridge University Press, 2008. Citado na página 68.

OHM, J.-R.; SCHAAR, M. V. D.; WOODS, J. W. Interframe wavelet coding-motion picture representation for universal scalability. *Signal Processing: Image Communication*, Elsevier, v. 19, n. 9, p. 877–908, 2004. Citado na página 25.

OLIVEIRA, P. et al. JPEG quantisation requires bit-shifts only. *Electronics Letters, IET*, v. 53, n. 9, p. 588–590, 2017. Citado 2 vezes nas páginas 41 e 106.

OLIVEIRA, P. A. M. et al. A discrete Tchebichef transform approximation for image and video coding. *IEEE Signal Processing Letters, IEEE*, v. 22, n. 8, p. 1137–1141, 2015. Citado 3 vezes nas páginas 24, 25 e 32.

OLIVEIRA, P. A. M. et al. Low-complexity image and video coding based on an approximate discrete Tchebichef transform. *IEEE Transactions on Circuits and Systems for Video Technology*. DOI: 10.1109/TCSVT.2016.2515378, 2016. Citado 9 vezes nas páginas 24, 25, 32, 33, 34, 38, 42, 45 e 51.

OPPENHEIM, A.; SCHAFER, R. *Discrete-Time Signal Processing*. [S.l.]: Pearson, 3a. edição, 2010. Citado 2 vezes nas páginas 34 e 104.

ORUKLU, E.; SANIIE, J. 3E-1 real-time ultrasonic imaging system based on discrete cosine transform for NDE applications. In: *IEEE Ultrasonics Symposium*. [S.l.: s.n.], 2006. p. 428–431. Citado na página 26.

OUYANG, X. et al. Low complexity discrete Hartley transform precoded OFDM for peak power reduction. *Electronics letters, IET*, v. 48, n. 2, p. 90–91, 2012. Citado 3 vezes nas páginas 25, 94 e 95.

PAIK, C. H.; FOX, M. D. Fast Hartley transforms for image processing. *IEEE transactions on medical imaging*, v. 7, n. 2, p. 149, 1988. Citado 2 vezes nas páginas 25 e 94.

PAPITHA, J.; NANCY, G. M.; NEDUMARAN, D. Compression techniques on MR image—a comparative study. In: IEEE. *Communications and Signal Processing (ICCSP), 2013 International Conference on*. [S.l.], 2013. p. 367–371. Citado 2 vezes nas páginas 27 e 95.

PEARSON, K. LIII. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Taylor & Francis, v. 2, n. 11, p. 559–572, 1901. Citado na página 88.

Pixmeo SARL Company. *OsiriX DICOM Image Library*. <http://www.osirix-viewer.com/resources/dicom-image-library/>. Disponível em: <<http://www.osirix-viewer.com/resources/dicom-image-library/>>. Citado na página 106.

POTLURI, U. S. et al. Improved 8-point approximate DCT for image and video compression requiring only 14 additions. *IEEE Transactions on Circuits and Systems I: Regular Papers*, v. 61, p. 1727–1740, 2014. Citado 7 vezes nas páginas 21, 22, 23, 66, 72, 73 e 74.

POURAZAD, M. T. et al. HEVC: The new gold standard for video compression: How does HEVC compare with H.264/AVC? *IEEE Consumer Electronics Magazine*, v. 1, p. 36–46, 2012. ISSN 2162-2248. Citado na página 22.

PRATT, W. K. Generalized Wiener filtering computation techniques. *IEEE Transactions on Computers*, IEEE, v. 100, n. 7, p. 636–641, 1972. Citado na página 19.

PRATTIPATI, S. et al. A fast 8×8 integer Tchebichef transform and comparison with integer cosine transform for image compression. In: IEEE. *56th International Midwest Symposium on Circuits and Systems (MWSCAS)*. [S.l.], 2013. p. 1294–1297. Citado 3 vezes nas páginas 24, 32 e 46.

PRATTIPATI, S.; SWAMY, M.; MEHER, P. A variable quantization technique for image compression using integer Tchebichef transform. In: IEEE. *9th International Conference on Information, Communications and Signal Processing (ICICSP)*. [S.l.], 2013. p. 1–5. Citado 2 vezes nas páginas 24 e 32.

RAO, K. R.; HWANG, J. J. *Techniques and Standards for Image, Video, and Audio Coding*. [S.l.]: Prentice Hall New Jersey, 1996. v. 70. Citado 2 vezes nas páginas 19 e 77.

RAO, K. R.; YIP, P. *The Transform and Data Compression Handbook*. [S.l.]: CRC Press LLC, 2001. Citado 7 vezes nas páginas 19, 20, 21, 51, 52, 65 e 81.

RAO, K. R.; YIP, P. *Discrete Cosine Transform: Algorithms, Advantages, Applications*. [S.l.]: Academic press, 2014. Citado 2 vezes nas páginas 19 e 20.

RICHARDSON, I. *The H.264 Advanced Video Compression Standard*. 2. ed. [S.l.]: John Wiley and Sons, 2010. Citado na página 53.

ROCHE, C. A split-radix partial input/output fast Fourier transform algorithm. *IEEE Transactions on signal processing*, v. 40, n. 5, p. 1273–1276, 1992. Citado na página 23.

ROSS, D. A. et al. Incremental learning for robust visual tracking. *International Journal of Computer Vision*, Springer, v. 77, n. 1-3, p. 125–141, 2008. Citado 3 vezes nas páginas 28, 89 e 90.

Rubo Medical Imaging. *Rubo Medical Imaging Database*. <http://www.rubomedical.com/company.html>. 2018. Disponível em: <<http://www.rubomedical.com/company.html>>. Citado na página 106.

SALOMON, D.; MOTTA, G. *Handbook of data compression*. [S.l.]: Springer Science & Business Media, 2010. Citado na página 51.

- SANG, A. et al. 6d vector orthogonal transformation and its application in multiview video coding. *The Imaging Science Journal*, Taylor & Francis, v. 61, n. 4, p. 341–350, 2013. Citado na página 27.
- SAPONARA, S. Real-time and low-power processing of 3D direct/inverse discrete cosine transform for low-complexity video codec. *Journal of Real-Time Image Processing*, Springer, v. 7, n. 1, p. 43–53, 2012. Citado na página 25.
- SAPONARA, S. Real-time and low-power processing of 3D direct/inverse discrete cosine transform for low-complexity video codec. *Journal of Real-Time Image Processing*, Springer Berlin / Heidelberg, v. 7, p. 43–53, 2012. ISSN 1861-8200. Disponível em: <<http://dx.doi.org/10.1007/s11554-010-0174-5>>. Citado 3 vezes nas páginas 26, 28 e 114.
- SAWANT, S.; ADJEROH, D. A. Balanced multiple description coding for 3D DCT video. *IEEE Transactions on Broadcasting*, v. 57, 2011. Citado 5 vezes nas páginas 26, 28, 65, 66 e 114.
- SENAPATI, R. K.; PATI, U. C.; MAHAPATRA, K. K. Reduced memory, low complexity embedded image compression algorithm using hierarchical listless discrete Tchebichef transform. *IET Image Processing*, IET, v. 8, n. 4, p. 213–238, 2014. Citado na página 32.
- SERVAIS, M.; JAGER, G. de. Video compression using the three dimensional discrete cosine transform (3D-DCT). In: IEEE. *Proceedings of the 1997 South African Symposium on Communications and Signal Processing (COMSIG)*. [S.l.], 1997. p. 27–32. Citado 7 vezes nas páginas 20, 26, 28, 76, 80, 81 e 84.
- Signal and Image Processing Institute. *USC-SIPI Image Database*. <http://sipi.use.edu/database/>. 2018. Disponível em: <<http://sipi.use.edu/database/>>. Citado 4 vezes nas páginas 49, 59, 60 e 137.
- SINGH, N.; SINHA, A. Optical image encryption using Hartley transform and logistic map. *Optics Communications*, Elsevier, v. 282, n. 6, p. 1104–1109, 2009. Citado na página 94.
- SMEULDERS, A. W. et al. Visual tracking: An experimental survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 36, n. 7, p. 1442–1468, 2014. Citado 3 vezes nas páginas 28, 89 e 92.
- SONG, T.; LI, H. Local polar DCT features for image description. *IEEE Signal Processing Letters*, v. 20, n. 1, p. 59–62, 2013. Citado 2 vezes nas páginas 20 e 27.
- SORENSEN, H. et al. On computing the discrete Hartley transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. 33, n. 5, p. 1231–1238, 1985. Citado 2 vezes nas páginas 25 e 95.
- SUEHIRO, N.; HATERI, M. Fast algorithms for the DFT and other sinusoidal transforms. *IEEE Transactions on Acoustic, Signal, and Speech Processing*, v. 34, p. 642–644, 1986. Citado na página 21.
- SULLIVAN, G. J. et al. Overview of the high efficiency video coding (HEVC) standard. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 22, p. 1649–1668, 2012. ISSN 1051-8215. Citado na página 53.

SULLIVAN, G. J. et al. Overview of the high efficiency video coding (HEVC) standard. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 22, p. 1649–1668, 2012. Citado na página 58.

SUN, Q. et al. A robust edge detection method with sub-pixel accuracy. *Optik-International Journal for Light and Electron Optics*, Elsevier, v. 125, n. 14, p. 3449–3453, 2014. Citado na página 31.

SUNDER, R. S.; ESWARAN, C.; SRIRAAM, N. Performance evaluation of 3-D transforms for medical image compression. In: IEEE. *2005 IEEE International Conference on Electro Information Technology*. [S.l.], 2005. p. 6–pp. Citado 2 vezes nas páginas 27 e 95.

SUNDER, R. S.; ESWARAN, C.; SRIRAAM, N. Medical image compression using 3-D Hartley transform. *Computers in Biology and Medicine*, Elsevier, v. 36, n. 9, p. 958–973, 2006. Citado 10 vezes nas páginas 25, 27, 94, 95, 96, 105, 106, 107, 108 e 110.

SURESH, K.; SREENIVAS, T. V. Block convolution using discrete trigonometric transforms and discrete Fourier transform. *IEEE Signal Processing Letters*, IEEE, v. 15, p. 469–472, 2008. Citado na página 19.

TEAM, J. V. *Recommendation H.264 and ISO/IEC 14 496–10 AVC: Draft ITU-T recommendation and final draft international standard of joint video specification*. [S.l.], 2003. Citado 2 vezes nas páginas 21 e 81.

TEAM x264. *x264*. 2017. <videolan.org/developers/x264.html>. Citado 2 vezes nas páginas 49 e 56.

TRAN, T. D. The binDCT: Fast multiplierless approximation of the DCT. *IEEE Signal Processing Letters*, v. 7, n. 6, p. 141–144, 2000. Citado 2 vezes nas páginas 21 e 82.

TSENG, C.-C.; LEE, S.-L. Design of digital IIR integrator using discrete Hartley transform interpolation method. In: IEEE. *Proceedings of the 2009 IEEE International Symposium on Circuits and Systems*. [S.l.], 2009. p. 2181–2184. Citado 2 vezes nas páginas 25 e 94.

VETTERLI, M.; NUSSBAUMER, H. Simple FFT and DCT algorithms with reduced number of operations. *Signal Processing*, v. 6, p. 267–278, 1984. Citado na página 21.

VILLASENOR, J. D. Alternatives to the discrete cosine transform for irreversible tomographic image compression. *IEEE Transactions on Medical Imaging*, v. 12, n. 4, p. 803–811, 1993. Citado 2 vezes nas páginas 25 e 94.

WALLACE, G. K. The JPEG still picture compression standard. *IEEE Transactions on Consumer Electronics*, v. 38, p. xviii–xxxiv, 1992. ISSN 0098-3063. Citado 6 vezes nas páginas 20, 21, 49, 65, 77 e 81.

WANG, C.-L.; CHANG, C.-H. A DHT-based FFT/IFFT processor for VDSL transceivers. In: IEEE. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001. (ICASSP'01)*. [S.l.], 2001. v. 2, p. 1213–1216. Citado 3 vezes nas páginas 25, 94 e 95.

WANG, X.; WANG, S.; BI, D. Distributed visual-target-surveillance system in wireless sensor networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, IEEE, v. 39, n. 5, p. 1134–1146, 2009. Citado na página 89.

WANG, Z. Fast algorithms for the discrete W transform and for the discrete Fourier transform. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. 32, p. 803–816, 1984. Citado na página 21.

WANG, Z. Pruning the fast discrete cosine transform. *IEEE Transactions on Communications*, v. 39, p. 640–643, 1991. ISSN 0090-6778. Citado 2 vezes nas páginas 22 e 23.

WANG, Z. et al. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, v. 13, p. 600–612, 2004. ISSN 1057-7149. Citado 4 vezes nas páginas 51, 52, 85 e 88.

XIANGWEN, D.; HEXIN, C.; ZHIJIE, Z. 4D-MDCT based adaptive prediction and compensation color video coding. In: IEEE. *Proceedings of 7th International Conference on Signal Processing*. [S.l.], 2004. v. 2, p. 1131–1134. Citado na página 27.

XIPH.ORG video test media (2014). Disponível em: <<https://media.xiph.org/video/derf/>>. Citado 2 vezes nas páginas 53 e 84.

XUE, M. et al. Automatic 4D facial expression recognition using DCT features. In: IEEE. *Winter Conference on Applications of Computer Vision*. [S.l.], 2015. p. 199–206. Citado na página 20.

YAKOVLEV, A.; KIM, S.; POON, A. Implantable biomedical devices: Wireless powering and communication. *IEEE Communications Magazine*, v. 50, n. 4, 2012. Citado na página 95.

YAP, P.-T.; PARAMESRAN, R. An efficient method for the computation of Legendre moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 27, n. 12, p. 1996–2002, 2005. Citado na página 31.

YOKOTANI, Y. et al. Lossless audio coding using the intMDCT and rounding error shaping. *IEEE Transactions on Audio, Speech, and Language Processing*, v. 14, n. 6, p. 2201–2211, 2006. Citado na página 21.

ZENG, Y.; BI, G.; LEYMAN, A. R. New polynomial transform algorithm for multidimensional DCT. *IEEE Transactions on Signal Processing*, v. 48, n. 10, p. 2814–2821, 2000. Citado 5 vezes nas páginas 20, 21, 27, 36 e 45.

ZENG, Y.; BI, G.; LEYMAN, A. R. New polynomial transform algorithm for multidimensional DCT. *IEEE Transactions on Signal Processing*, IEEE, v. 48, n. 10, p. 2814–2821, 2000. Citado na página 27.

Zenodo.org. *An MRI DICOM data set of the head of a normal male human aged 52*. <https://zenodo.org/record/16956>. Disponível em: <<https://zenodo.org/record/16956>>. Citado na página 106.

ZHANG, C.; LI, J. Compression of lumigraph with multiple reference frame (MRF) prediction and just-in-time rendering. In: IEEE. *Proceedings of Data Compression Conference*. [S.l.], 2000. p. 253–262. Citado na página 27.

ZHANG, H. et al. Symmetric image recognition by Tchebichef moment invariants. In: IEEE. *Image Processing (ICIP), 2010 17th IEEE International Conference on*. [S.l.], 2010. p. 2273–2276. Citado na página 32.

ZHANG, H. et al. Affine Legendre moment invariants for image watermarking robust to geometric distortions. *IEEE Transactions on Image Processing*, v. 20, n. 8, p. 2189–2199, 2011. Citado na página 31.

ZHAO, D.; LI, X.; CHEN, L. Optical image encryption with redefined fractional Hartley transform. *Optics Communications*, Elsevier, v. 281, n. 21, p. 5326–5329, 2008. Citado na página 94.

ZHOU, J. et al. FOFDM based on discrete cosine transform for intensity-modulated and direct-detected systems. *Journal of lightwave technology*, IEEE, v. 34, n. 16, p. 3717–3725, 2016. Citado na página 19.

ZHOU, T. et al. Robust visual tracking via efficient manifold ranking with low-dimensional compressive features. *Pattern Recognition*, Elsevier, v. 48, n. 8, p. 2459–2473, 2015. Citado na página 90.

ZHU, H. et al. General form for obtaining discrete orthogonal moments. *IET Image Processing*, IET, v. 4, n. 5, p. 335–352, 2010. Citado na página 24.

APÊNDICE A – IMAGENS EMPREGADAS NA SIMULAÇÃO DE COMPRESSÃO JPEG

A seguir, são apresentadas todas as imagens empregadas dos algoritmos de simulação de compressão de imagem estática.



Fonte: (Signal and Image Processing Institute, 2018)

APÊNDICE B – TRABALHOS PRODUZIDOS E PUBLICADOS

As seguintes contribuições bibliográficas acadêmicas e de inovação foram derivadas da presente tese de doutorado:

- **Manuscritos publicados:**

1. COUTINHO, V. A.; CINTRA, R. J.; BAYER, F. M. Low-complexity multidimensional DCT approximations for high-order tensor data decorrelation. *IEEE Transactions on Image Processing*, v. 26, n. 5, p. 2296–2310, 2017.
2. COUTINHO, V. A.; CINTRA, R. J.; BAYER, F. M.; OLIVEIRA, P. A. M.; OLIVEIRA, R. S.; MADANAYAKE, A. Pruned discrete Tchebichef transform approximation for image compression. *Circuits, Systems, and Signal Processing*, Springer, p. 1–21, 2018.

- **Patentes depositadas:**

1. COUTINHO, V. A.; CINTRA, R. J.; BAYER, F. M. Codificador tridimensional de imagem e vídeo de baixo consumo energético. *Patente depositada junto ao INPI*, n. BR 10 2017 023911 0, 2017.
2. COUTINHO, V. A.; CINTRA, R. J.; BAYER, F. M. Sistema de rastreamento visual baseado em transformada tridimensional de baixa complexidade. *Patente depositada junto ao INPI*, n. BR 10 2017 024085 1, 2017.

- **Manuscrito em preparação:**

1. COUTINHO, V. A.; CINTRA, R. J.; BAYER, F. M. Low-complexity three-dimensional discrete Hartley transform approximation for medical image compression.

Adicionalmente, os seguintes trabalhos acadêmicos e produções bibliográficas correlatas foram realizados em paralelo durante o período de doutorado:

- **Estágio de doutorado sanduíche.** Realizado em *The University of Akron*, Akron, Ohio, EUA, com duração de 10 meses. O autor deste trabalho atuou junto ao *Advanced Signal Processing Circuits Lab* sob a orientação do Prof. Dr. Arjuna Madanayake, trabalhando em implementações em VLSI e FPGA nas áreas de *beamforming*, filtros bidimensionais, filtros gaussianos, aproximações de baixa complexidade para a DFT e outros sistemas lineares.

- **Manuscritos publicados:**

1. COUTINHO, V. A.; ARIYARATHNA, V.; COELHO, D. F. G.; CINTRA, R. J.; MADANAYAKE, A. An 8-beam 2.4 GHz digital array receiver based on a fast multiplierless spatial DFT approximation. In: IEEE. *2018 IEEE/MTT-S International Microwave Symposium-IMS*. [S.l.], 2018. p. 1538–1541.
2. ARIYARATHNA, V.; COUTINHO, V. A.; PULIPATI, S.; MADANAYAKE, A.; WIJESEKARA, R. T.; EDUSSOORIYA, C. U. S.; BRUTON, L. T.; GUNARATNE, T. K.; CINTRA, R. J. Real-time 2-D FIR trapezoidal digital filters for 2.4 GHz aperture receiver applications. In: IEEE. *2018 Moratuwa Engineering Research Conference (MERCOn)*. [S.l.], 2018. p. 350–355. **O artigo acima foi vencedor do prêmio ‘best paper award’ na categoria ‘Image, Vision, and Signal Processing’.**

- **Capítulo de livro em preparação:**

1. COELHO, D. F. G.; COUTINHO, V. A.; SILVEIRA, T.; CINTRA, R. J.; BAYER, F. M.; MADANAYAKE, A.; DIMITROV, V. Discrete transform approximations for video coding. In: MARTINA, M. (Ed.). *VLSI Architectures for Future Video Coding*. [S.l.]: The Institution of Engineering and Technology (IET).

- **Manuscritos em preparação:**

1. MARTINEZ, P. A.; COUTINHO, V. A.; CINTRA, R. J.; LEITE, A.; BAYER, F. M.; MADANAYAKE, A. Gaussian kernel approximations require only bit-shifts.

2. COUTINHO, V. A.; ARIYARATHNA, V.; COELHO, D. F. G.; CINTRA, R. J.; MADANAYAKE, A. A 16-beam 2.4 GHz digital array receiver based on fast multiplierless spatial DFT approximations.