



República Federativa do Brasil
Ministério da Economia
Instituto Nacional da Propriedade Industrial

(21) BR 102017023911-0 A2



(22) Data do Depósito: 07/11/2017

(43) Data da Publicação Nacional: 04/06/2019

(54) **Título:** CODIFICADOR TRIDIMENSIONAL DE IMAGEM E VÍDEO DE BAIXO CONSUMO ENERGÉTICO

(51) **Int. Cl.:** H04N 19/625; H04N 19/00; G06T 9/00.

(52) **CPC:** H04N 19/625; H04N 19/00; G06T 9/007; G06T 9/00.

(71) **Depositante(es):** UNIVERSIDADE FEDERAL DE PERNAMBUCO; UNIVERSIDADE FEDERAL DE SANTA MARIA.

(72) **Inventor(es):** VÍTOR DE ANDRADE COUTINHO; RENATO JOSÉ DE SOBRAL CINTRA; FÁBIO MARIANO BAYER.

(57) **Resumo:** Esta invenção trata de um esquema de compressão de imagens ou vídeos nos mais variados formatos de dados tridimensionais, apresentando baixo custo computacional e alto desempenho de compactação. O método pode ser aplicado para imagens dinâmicas como vídeo e imagens estáticas tridimensionais. Este esquema se baseia em uma nova transformação tridimensional baseada em uma matriz de transformação associada a uma aproximação para a DCT de baixo custo computacional. A nova transformada tridimensional apresentada evita a necessidade de operações em ponto flutuante durante a etapa de compactação e descompactação dos dados processados. O bloco da transformada usualmente implementada pela transformada discreta do cosseno tridimensional (3D DCT) é substituído pela transformada tridimensional de baixa complexidade apresentada, que possui alta performance de decorrelação de vídeos e outros sinais classificados como processo estacionário Markov do tipo I. Além da substituição do bloco da 3D DCT por uma outra transformada, o bloco de quantização é modificado de modo a tornar a complexidade da transformada apresentada baixa. Além disso, o presente invento apresenta grande eficiência energética por fazer uso de uma matriz com coeficientes apenas em 0 e ± 1 . Isso implica em eliminação da complexidade multiplicativa da etapa de processamento do sinal tridimensional e torna desnecessária a utilização (...).



CODIFICADOR TRIDIMENSIONAL DE IMAGEM E VÍDEO DE BAIXO CONSUMO ENERGÉTICO

CAMPO DE APLICAÇÃO

[001] A presente invenção tem aplicabilidade em compressão de vídeos digitais e imagens tridimensionais com alto grau de correlação espacial e temporal. O dispositivo se baseia em transformada de bloco tridimensional. A transformada discreta tridimensional do cosseno (3D DCT) pode ser aplicada para compressão de vídeos explorando a correlação entre *pixels* nas dimensões espaciais e temporal [33]. Entretanto, tal ferramenta matemática apresenta as desvantagens de possuir custo computacional elevado devido à necessidade de operações de multiplicação e de aritmética de ponto flutuante, que demandam alto consumo de energia e recursos de *hardware*. Dessa forma, o usual sistema baseado na 3D DCT não é adequado para aplicações embarcadas com restrições críticas de consumo energético e capacidade computacional. A presente invenção se baseia na substituição do bloco da 3D DCT por um novo esquema de baixa complexidade para computar a transformada tridimensional de bloco com alta capacidade de compactação de energia.

ANTECEDENTE DA INVENÇÃO

[002] Imagens e vídeos digitais geralmente são representados por *arrays* multidimensionais de *pixels* que, geralmente, são correlacionados e podem ser modelados como processos markovianos de primeira ordem com elevado coeficiente de correlação [34]. Para tais tipos de dados, a transformada discreta do cosseno (DCT) se comporta como uma aproximação para a transformada de Karhunen-Loève (KLT) [35], que é a transformada ótima para decorrelação de dados [36]. A vantagem do uso da DCT se dá pelo fato desta apresentar núcleo de transformação fixo, enquanto o da KLT depende do comportamento estatístico dos dados de entrada. Em consequência de ser mais prático o desenvolvimento de algoritmos rápidos que dependem apenas do comprimento da transformada [37], a DCT é aplicada em diversas tecnologias de

compressão de dados [38].

[003] A DCT transforma um sinal discreto $\mathbf{x} = [x[0] \ x[1] \ \dots \ x[N-1]]^\top$ no sinal de saída $\mathbf{X} = [X[0] \ X[1] \ \dots \ X[N-1]]^\top$, de acordo com a seguinte relação [39]:

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

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

em que

$$\alpha_N[k] \triangleq \sqrt{\frac{1}{N}} \cdot \begin{cases} 1, & \text{if } k = 0, \\ \sqrt{2}, & \text{otherwise.} \end{cases} \quad (2)$$

[004] Uma imagem pode ser representada como uma matriz—*array* de duas dimensões—de *pixels*, em que cada elemento da matriz carrega um valor inteiro que representa intensidade de cor para cada uma das cores vermelha, verde e azul (RGB). Considerando a representação usual de 8 bits, tais valores de intensidade variam entre 0 a 255. Em escala de cinza, é considerado apenas um valor para cada elemento da matriz. Similarmente, um vídeo pode ser representado por uma sucessão de imagens (*frames* ou quadros). Em padrões de codificação de vídeo, tais como MPEG [40, 41], H.261 [42] H.263 [43], H.264 [44], HEVC [45], a DCT bidimensional (2D DCT) é aplicada a cada quadro de vídeo para decorrelação espacial. Seja \mathbf{A} um sinal bidimensional, representado por uma matriz de tamanho $N_1 \times N_2$, cujas entradas são dadas por $a[n_1, n_2]$, para $n_i = 1, 2, \dots, N_i - 1$, e $i = 1, 2$. As entradas do sinal transformado \mathbf{B} são computadas de acordo com: [46]

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

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

[005] Apesar de ser largamente empregada, a 2D DCT não leva em conta a correlação

entre *pixels* de quadros consecutivos. Em geral, padrões de codificação de vídeo aplicam algoritmos de estimação de movimento para explorar a correlação espacial [47], que são consideravelmente custosos computacionalmente [48]. Uma alternativa para evitar tais algoritmos é aplicar codificação de vídeo *entre-quadros*, que aplica a 3D DCT a consecutivos quadros de vídeo [49, 50, 51, 52, 53, 54, 55, 48, 56, 33, 57]. Tal método divide quadros sucessivos em “cubos” menores de *pixels*, os quais são transformados [48]. Os esquemas usuais de codificação e decodificação são mostrados nas Figura 1 e Figura 2, respectivamente. A 3D DCT de um sinal \mathcal{T} com entradas $t[n_1, n_2, n_3]$, $n_i = 0, 1, \dots, N_i - 1$, para $i = 1, 2, 3$, é dado pelo sinal \mathcal{Y} , cujas entradas são dadas por [58, 59]:

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

$$k_i = 1, 2, \dots, N_i - 1, \quad i = 1, 2, 3. \quad (4)$$

[006] Vetores e matrizes podem ser modelados como *tensores* de primeira e segunda ordem, respectivamente [60]. De forma análoga, um sinal 3D pode ser entendido como um tensor de terceira ordem [61, 62, 59]. Um tensor de ordem R -ésima é simplesmente um *array* que requer R índices. Seja $\mathcal{A} \in \mathbb{C}^{N_1 \times N_2 \times \dots \times N_R}$ um tensor de R -ésima ordem, cujas entradas são dadas por $a[n_1, n_2, \dots, n_R]$, em que \mathbb{C} é o conjunto dos números complexos e $n_i = 0, 1, \dots, N_i - 1$, para $i = 1, 2, \dots, R$. O *produto em i -ésimo modo* de tensor \mathcal{A} pela matriz $\mathbf{M} \in \mathbb{C}^{H \times N_i}$ [63, p. xxxv], denotado por $\mathcal{A} \times_i \mathbf{M}$, é definido como o tensor $\mathcal{B} \in \mathbb{C}^{N_1 \times N_2 \times \dots \times N_{i-1} \times H \times N_{i+1} \times \dots \times N_R}$, cujas entradas são dadas por:

$$b[n_1, \dots, n_{i-1}, h, n_{i+1}, \dots, n_R] \triangleq \sum_{n_i=0}^{N_i-1} a[n_1, \dots, n_i, \dots, n_R] \cdot m[h, n_i], \quad (5)$$

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 generaliza o produto matricial, viabilizando operar *arrays* de alta ordem com matrizes.

A DCT de múltiplas dimensões pode ser expressa em termos de produtos em i -ésimo modo pela matriz de transformação da DCT. A DCT unidimensional em (1) pode ser expressa por $\mathbf{X} = \mathbf{x} \times_1 \mathbf{C}_N$, em que \mathbf{C}_N é a matriz da DCT, cujas entradas são dadas por:

$$c_N[k, n] = \alpha_N[k] \cdot \cos\left(\frac{\pi(2n+1)k}{2N}\right), \quad (6)$$

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

Analogamente, A 2D DCT em (3) equivale a $\mathbf{B} = \mathbf{A} \times_1 \mathbf{C}_{N_1} \times_2 \mathbf{C}_{N_2}$ e a 3D DCT em (4) é pode ser escrita na forma [59]:

$$\mathcal{Y} = \mathcal{T} \times_1 \mathbf{C}_{N_1} \times_2 \mathbf{C}_{N_2} \times_3 \mathbf{C}_{N_3}. \quad (7)$$

Uma vez que a matriz da DCT é ortogonal, ou seja, $\mathbf{C}_N^{-1} = \mathbf{C}_N^\top$ as transformadas inversas para a 1D DCT, 2D DCT e 3D DCT são dadas respectivamente por:

$$\mathbf{x} = \mathbf{X} \times_1 \mathbf{C}_N^\top \quad (8)$$

$$\mathbf{A} = \mathbf{B} \times_1 \mathbf{C}_{N_1}^\top \times_2 \mathbf{C}_{N_2}^\top \quad (9)$$

$$\mathcal{T} = \mathcal{Y} \times_1 \mathbf{C}_{N_1}^\top \times_2 \mathbf{C}_{N_2}^\top \times_3 \mathbf{C}_{N_3}^\top. \quad (10)$$

[007] Após o estágio da transformação 3D no esquema usual de compressão, mostrado na Figura 1, os coeficiente obtidos são quantizados por um *volume de quantização* e reorganizados de acordo com esquema *zig-zag* 3D [56]. Seja $\mathcal{Q} \in \mathbb{C}^{N_1 \times N_2 \times N_3}$ o volume de quantização, tensor de terceira ordem cujas entradas são dadas por $q[k_1, k_2, k_3]$, para $k_1, k_2, k_3 = 0, 1, \dots, N_i - 1, i = 1, 2, 3$. A etapa de quantização realiza a seguinte operação: [33, 48]:

$$\tilde{y}[k_1, k_2, k_3] = \text{round}\left(\frac{y[k_1, k_2, k_3]}{q[k_1, k_2, k_3]}\right), \quad (11)$$

$$k_1, k_2, k_3 = 0, 1, \dots, N_i - 1, \quad i = 1, 2, 3,$$

onde $y[k_1, k_2, k_3]$ são os coeficientes no domínio da frequência de acordo com (4). Os

coeficientes do volume de quantização devem ser projetados para preservar ou descartar informação de vídeo no domínio da frequência de acordo com algum critério de *trade-off* entre qualidade visual e taxa de compressão. O processo de desquantização é definido por $\hat{y}[k_1, k_2, k_3] = \tilde{y}[k_1, k_2, k_3] \cdot q[k_1, k_2, k_3]$, onde $\tilde{y}[k_1, k_2, k_3]$ é uma estimativa de $y[k_1, k_2, k_3]$ [64].

[008] Algoritmos multidimensionais para a DCT frequentemente aplicam a propriedade de *separação de núcleo*, que permite o uso de algoritmos para computar transformadas de dimensões maiores [37, 65]. Esse método é frequentemente denominado *decomposição linha-coluna* [66]. Tendo em vista a grande demanda de aplicações, vários circuitos esquemáticos rápidos para a 1D DCT foram propostos [67, 68, 69, 70, 71, 72]. De fato, a complexidade multiplicativa mínima teórica prevista por Heideman [73] foi atingida pelo algoritmo de Loeffler para o comprimento $N = 8$ [71]. Uma vez que se trata de um campo de pesquisa já bem desenvolvido, é improvável que novos algoritmos apresentem melhorias significativas em termos de complexidade computacional. Neste contexto, diferentes métodos aproximados foram visados para diminuir o custo do cômputo da DCT, tal como a DCT inteira [74, 36, 75, 76, 77], DCTs aproximadas [78, 79, 80, 81, 82], algoritmos podados para a DCT [83, 84, 85], assim como técnicas combinadas [86, 87, 88, 89]. Em particular, DCTs aproximadas são transformadas que apresentam custo computacional muito baixo (em geral o custo multiplicativo é nulo), enquanto preservam características importantes da transformação original, como compactação de energia e capacidade de decorrelação. Diferentemente da DCT original—também chamada de *DCT exata*— as aproximações não estão sujeitas aos limites multiplicativos teóricos.

[009] Diversas aproximações para a DCT vem sendo propostas nos últimos anos. Neste contexto, a 2D DCT aproximada vem sendo aplicada com sucesso em aplicações de compressão de imagem, obtendo performance competitiva a um custo computacional consideravelmente menor [90, 87, 88]. Apesar de vasto espectro de aplicações, o projeto de aproximações para a DCT de múltiplas dimensões representam um campo industrial inexplorado.

SUMÁRIO DA INVENÇÃO

[010] A presente invenção propõe um dispositivo para a compressão de imagens 3D ou vídeos com baixo custo computacional e notória capacidade de codificação. O esquema se baseia em aplicar produto em i -ésimo de tensores de alta ordem por uma matriz aproximada para DCT proposta em [81], derivando uma 3D DCT de tamanho $8 \times 8 \times 8$. Tal transformada possui a menor complexidade aritmética da literatura, com apenas 14 adições em 1D e não necessita de operações com ponto flutuantes. A aproximação para a 3D DCT proposta requer apenas 2688 adições, enquanto a 3D DCT exata necessita de 2112 multiplicações e 5568 adições para o mesmo tamanho de transformada, derivada a partir do algoritmo de Loeffler [71].

[011] Os bloco da transformadas direta e inversa no esquema de compactação e descompactação de imagem e video, usualmente implementado pela 3D DCT e 3D DCT inversa (3D IDCT), respectivamente, são substituído pelos métodos propostos nesta invenção. O bloco de quantização 3D e desquantização 3D também são modificados de modo a explorar a baixa complexidade aritmética da transformação tridimensional. A invenção propões um método industrializável relevante, diferindo dos métodos já existentes, se tratando de esquemas de codificação de imagem cujo bloco de transformada 3D direta e inversa dos esquemas gerais das Figuras **1** e **2** são substituídas por aproximações para a 3D DCT e 3D IDCT, respectivamente, que possuem menor complexidade computacional que os métodos já existentes, resultando nos esquemas de compressão e descompressão propostos mostrados nas Figuras **3** e **4**. Os blocos de quantização e desquantização são modificados de modo a explorar melhor as propriedades de baixa complexidade das transformações tridimensionais propostas.

BREVE DESCRIÇÃO DAS FIGURAS

[012] A Figura **1** apresenta esquema usual de compressão de vídeo através de divisão em blocos 3D, transformada de bloco 3D, quantização 3D e codificação de entropia.

[013] A Figura **2** apresenta esquema usual de descompressão de vídeo comprimido com o esquema da Figura **1**.

[014] A Figura **3** apresenta o esquema proposto de compressão de vídeo através de divisão em blocos 3D, transformada aproximada proposta para a 3D DCT, quantização modificada 3D e codificação de entropia.

[015] A Figura 4 apresenta o esquema proposto de descompressão de vídeo comprimido com o esquema da Figura 3.

[016] A Figura 5 mostra o algoritmo rápido referente para o calculo da DCT unidimensional que é empregado para computar a 3D DCT de acordo com (16) e (17).

[017] A Figura 6 mostra o algoritmo rápido referente para o calculo da DCT inversa unidimensional que é empregado para computar a 3D DCT inversa de acordo com (25) e (17).

DESCRIÇÃO DETALHADA

[018] O novo esquema proposto para compressão de vídeos é apresentado a seguir. O esquema se baseia em uma nova transformada tridimensional aqui proposta. A presente invenção difere significativamente de esquemas já propostos em trabalhos relacionados.

[019] A invenção proposta em BRPI05060842A [91] sugere transformadas com elementos pertencentes ao conjunto dos inteiros diádicos [92]. Apesar de reduzir a complexidade computacional em relação a DCT exata, ainda há a necessidade de multiplicações não triviais. Além disso, é proposta uma transformada bidimensional. A presente invenção sugere transformadas para codificação tridimensional baseadas em matrizes retangulares podadas com elementos do conjunto $\{0,1\}$.

[020] O método sugerido em US7336720B2 [93] propões um sistema de codificação que emprega ambas a 2D DCT e 3D DCT exatas, diferindo da presente invenção, que propõe um esquema baseado em uma aproximação para a 3D DCT.

[021] O método proposto em US3678582657B2 [94] sugere um esquema de transformação multidimensional de uma pluralidade de blocos, em que a DCT multidimensional exata é empregada, diferindo, portanto, da presente invenção.

[022] A invenção reivindicada em US20120057777A1 [95] propõe um codificador baseado na 3D DCT exata, apresentando custo computacional mais elevado que o esquema proposto na presente invenção, em que um esquema baseado em uma 3D DCT aproximada de baixo custo computacional é proposto.

[023] O método proposto em US5126962A [96] sugere um sistema de processar a 2D DCT e 3D DCT exatas por meio da DCT unidimensional exata, fazendo uso da decomposição linha coluna, diferindo da presente invenção, que propõe um esquema de

codificação baseado em uma 3D DCT aproximada.

[024] O esquema patenteado em US6535629B2 [97] propõe um sistema de codificação de imagem 3D que faz uso da 3D DCT exata e quantização, diferindo da presente invenção, que faz uso da 3D DCT aproximada proposta e um bloco modificado de quantização.

[025] O método sugerido em US7224845B1 [98] propõe um esquema de compressão de imagens multiespectral em que uma imagem 2D virtual é gerada através de um mapeamento bijetivo e então comprimida através da KLT, diferindo, portanto, da presente invenção, que aplica a 3D DCT aproximada proposta.

[026] O esquema proposto em US20120251014A1 [99] representa um sistema de processamento de dados multidimensionais que se baseia em transformadas multidimensionais ortogonais. A presente invenção difere por propor uma aproximação ortogonal para a 3D DCT baseado em circuito esquemático para a DCT aproximada.

[027] A invenção reivindicada em RU2375838C1 [100] propõe esquema de codificação e decodificação de imagens baseado na 3D DCT exata, diferindo da presente invenção, que propõe método baseado na 3D DCT aproximada proposta.

[028] O esquema proposto se baseia em aplicar álgebra tensorial combinado com a matriz aproximada DCT proposta em [81] e obter uma aproximação tridimensional para a DCT de tamanho $8 \times 8 \times 8$. Tal transformada tridimensional apresentam custo multiplicativo nulo e é aplicada aos esquemas de compressão e descompressão de vídeo baseado em transformada 3D, resultando nos esquemáticos mostrados nas Figura 3 e Figura 4. Um novo design para o bloco de quantização e desquantização multidimensionais são propostos de modo a explorar as propriedades de baixa complexidade das transformadas propostas.

[029] Seja a matriz de transformação de baixa complexidade

$$\mathbf{T} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (12)$$

que corresponde a uma aproximação para a DCT dada por $\hat{\mathbf{C}} = \mathbf{S} \cdot \mathbf{T}$, em que $\mathbf{S} = \text{diag} \left(\frac{1}{\sqrt{8}}, \frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{2}} \right)$. Definimos uma aproximação 3D de um sinal

$\mathcal{T} \in \mathbb{C}^{N_1 \times N_2 \times N_3}$, representado por tensor de terceira ordem, o sinal tridimensional de saída dado por

$$\begin{aligned} \mathcal{Y} &= \mathcal{T} \times_1 \hat{\mathbf{C}} \times_2 \hat{\mathbf{C}} \times_3 \hat{\mathbf{C}} \\ &= \mathcal{T} \times_1 (\mathbf{S} \cdot \mathbf{T}) \times_2 (\mathbf{S} \cdot \mathbf{T}) \times_3 (\mathbf{S} \cdot \mathbf{T}). \end{aligned} \quad (13)$$

Considerando propriedades algébricas do produto em i -ésimo modo [61, 60], verifica-se que a propriedade da separabilidade é preservada. Além disso, pode-se reorganizar (13) da seguinte forma:

$$\mathcal{Y} = \mathcal{T} \times_1 \mathbf{T} \times_2 \mathbf{T} \times_3 \mathbf{T} \times_1 \mathbf{S} \times_2 \mathbf{S} \times_3 \mathbf{S}. \quad (14)$$

Sendo $\hat{\mathbf{C}}$ uma matriz de transformação ortogonal e $\mathbf{S} = \mathbf{S}^\top$ uma matriz diagonal, tem-se que $\hat{\mathbf{C}}^{-1} = \hat{\mathbf{C}}^\top = \mathbf{T}^\top \cdot \mathbf{S}$. Conseqüentemente, a transformação inversa para (14) é dada por:

$$\begin{aligned} \mathcal{T} &= \mathcal{Y} \times_1 \hat{\mathbf{C}}_N^{-1} \times_2 \hat{\mathbf{C}}_N^{-1} \times_3 \hat{\mathbf{C}}_N^{-1} \\ &= \mathcal{Y} \times_1 (\mathbf{S} \cdot \mathbf{T})^\top \times_2 (\mathbf{S} \cdot \mathbf{T})^\top \times_3 (\mathbf{S} \cdot \mathbf{T})^\top \\ &= \mathcal{Y} \times_1 \mathbf{S} \times_2 \mathbf{S} \times_3 \mathbf{S} \times_1 \mathbf{T}^\top \times_2 \mathbf{T}^\top \times_3 \mathbf{T}^\top. \end{aligned} \quad (15)$$

[030] A 3D DCT aproximada proposta pode ser computada em duas etapas separadas: a primeira envolvendo os produtos em i -ésimo modo com as matrizes de baixa complexidade \mathbf{T} e na etapa subsequente, pelas matrizes \mathbf{S} , que apresentam entradas dada por números irracionais. Sugerimos um método de embutir tais valores irracionais na etapa subsequente de quantização, de modo que a matriz \mathbf{S} não introduza complexidade adicional. Definimos o sinal seguinte tensor de terceira ordem intermediário:

$$\mathcal{A} = \mathcal{T} \times_1 \mathbf{T} \times_2 \mathbf{T} \times_3 \mathbf{T}, \quad (16)$$

Para cada produto em i -ésimo modo e mantendo os demais índices das outras dimensões fixos, temos o equivalente a um produto matricial unidimensional. Sejam $t[n_1, n_2, n_3]$ e $a[k_1, n_2, n_3]$ as entradas dos sinais \mathcal{T} e \mathcal{A} , respectivamente, para $n_1, k_1 = 0, 1, \dots, 7$,

fixando n_2 e n_3 . Então, $\mathcal{A} = \mathcal{T} \times_1 \mathbf{T}$ pode ser representado pelo produto

$$\begin{bmatrix} X[0] \\ X[1] \\ \vdots \\ X[7] \end{bmatrix} = \mathbf{T} \cdot \begin{bmatrix} x[0] \\ x[1] \\ \vdots \\ x[7] \end{bmatrix}, \quad (17)$$

$$X[k_1] \triangleq a[k_1, n_2, n_3] \quad x[n_1] \triangleq t[n_1, n_2, n_3].$$

O produto (17) pode ser calculado pelo circuito esquemático rápido mostrado na Figura 5. Ao variar, os valores de n_2 e n_3 , é necessário $N^2 = 8^2$ cálculos da expressão (17) e, portanto, tal quantidade de aplicações do circuito esquemático rápido da Figura 5. Para os demais produtos \times_2 e \times_3 , o desenvolvimento é análogo, totalizando $3 \cdot N^2 = 3 \cdot 8^2 = 192$ aplicação do circuito esquemático rápido. Sendo assim, a expressão (16) aproximada proposta pode ser computada através do circuito esquemático rápido unidimensional apresentado na Figura 5, que demanda 14 adições individualmente, totalizando $192 \cdot 14 = 2688$ adições para o cálculo da 3D DCT aproximada completa.

[031] Para obter a 3D DCT aproximada proposta (14), é necessário realizar realizar os produtos em i -ésimo modo do sinal \mathbf{A} pela matriz diagonal \mathbf{S} de acordo com:

$$\mathcal{Y} = \mathbf{A} \times_1 \mathbf{S} \times_2 \mathbf{S} \times_3 \mathbf{S}. \quad (18)$$

Como \mathbf{S} é uma matriz diagonal, é possível demonstrar que:

$$y[k_1, k_2, k_3] = a[k_1, k_2, k_3] \cdot d_{k_1} \cdot d_{k_2} \cdot d_{k_3}, \quad (19)$$

em que $a[k_1, k_2, k_3]$ são as entradas do sinal \mathcal{A} dados em (16) e d_k é o k -ésimo elemento diagonal de \mathbf{S} . O cálculo da expressão (19) pode ser embutido na etapa subsequente de quantização. De (19) em (11), com $N_1 = N_2 = N_3 = 8$, obtemos:

$$\tilde{y}[k_1, k_2, k_3] = \text{round} \left(\frac{a[k_1, k_2, k_3] \cdot d_{k_1} \cdot d_{k_2} \cdot d_{k_3}}{q[k_1, k_2, k_3]} \right), \quad (20)$$

$$k_1, k_2, k_3 = 0, 1, \dots, 7.$$

[032] Propomos um novo volume de quantização $\mathcal{Q}^* \in \mathbb{C}^{8 \times 8 \times 8}$, cujas entradas são dadas por:

$$q^*[k_1, k_2, k_3] \triangleq \frac{q[k_1, k_2, k_3]}{d_{k_1} \cdot d_{k_2} \cdot d_{k_3}}. \quad (21)$$

Substituindo (21) em (20), obtemos o volume de quantização modificado como se segue:

$$\tilde{y}[k_1, k_2, k_3] = \text{round} \left(\frac{a[k_1, k_2, k_3]}{q^*[k_1, k_2, k_3]} \right), \quad (22)$$

$$k_1, k_2, k_3 = 0, 1, \dots, 7.$$

Apenas o sinal \mathcal{A} é necessário ser computado. Tal sinal é aplicado como entrada ao volume de quantização modificado (22). Uma vez que o cômputo de \mathcal{A} requer apenas a transformação 3D de baixa complexidade (16), o excedente computacional imposto pela matriz diagonal \mathbf{S} é descartado. O procedimento de quantização inversa modificado é obtido de forma análoga, resultando num tensor $\hat{\mathcal{A}}$, cujas entradas são dadas por:

$$\hat{a}[k_1, k_2, k_3] = \tilde{y}[k_1, k_2, k_3] \cdot q^\dagger[k_1, k_2, k_3], \quad (23)$$

em que

$$q^\dagger[k_1, k_2, k_3] \triangleq q[k_1, k_2, k_3] \cdot d_{k_1} \cdot d_{k_2} \cdot d_{k_3} \quad (24)$$

são as entradas do volume de quantização inverso modificado \mathcal{Q}^\dagger . De maneira análoga ao caso direto, na presente invenção, apenas as matrizes de baixa complexidade são empregadas na etapa inversa, de acordo com

$$\hat{\mathcal{T}} = \hat{\mathcal{A}} \times_1 \mathbf{T}^\top \times_2 \mathbf{T}^\top \times_3 \mathbf{T}^\top, \quad (25)$$

em que $\hat{\mathcal{T}}$ é um bloco recuperado do vídeo comprimido. A transformação inversa também pode ser aplicada com a decomposição-linha coluna, fazendo uso do algoritmo rápido 1D mostrado na Figura 6, analogamente ao caso direto.

Referências

- [33] M. C. Lee, Raymond K. W. Chan, and D. A. Adjeroh. Quantization of 3D-DCT coefficients and scan order for video compression. *Journal of Visual Communication and Image Representation*, 8(4):405–422, 1997.
- [34] K. R. Rao and P. Yip. *The Transform and Data Compression Handbook*. CRC Press LLC, 2001.
- [35] R. J. Clarke. Relation between the Karhunen-Loève and cosine transforms. *IEEE Proceedings F Communications, Radar and Signal Processing*, 128:359–360, 1981.
- [36] V. Britanak, P. Yip, and K. R. Rao. *Discrete Cosine and Sine Transforms*. Academic Press, 2007.
- [37] R. E. Blahut. *Fast Algorithms for Signal Processing*. Cambridge University Press, 2010.
- [38] K. R. Rao and P. Yip. *Discrete Cosine Transform: Algorithms, Advantages, Applications*. Academic Press, San Diego, CA, 1990.
- [39] N. Ahmed, T. Natarajan, and K. R. Rao. Discrete cosine transform. *IEEE Transactions on Computers*, C-23:90–93, 1974.
- [40] G. K. Wallace. The JPEG still picture compression standard. *IEEE Transactions on Consumer Electronics*, 38:xviii–xxxiv, 1992.
- [41] International Organisation for Standardisation. Generic coding of moving pictures and associated audio information – Part 2: Video. ISO/IEC JTC1/SC29/WG11 - coding of moving pictures and audio, ISO, 1994.
- [42] International Telecommunication Union. ITU-T recommendation H.261 version 1: Video codec for audiovisual services at $p \times 64$ kbits. Technical report, ITU-T, 1990.
- [43] International Telecommunication Union. ITU-T recommendation H.263 version 1: Video coding for low bit rate communication. Technical report, ITU-T, 1995.

- [44] Joint Video Team. Recommendation H.264 and ISO/IEC 14 496–10 AVC: Draft ITU-T recommendation and final draft international standard of joint video specification. Technical report, ITU-T, 2003.
- [45] International Telecommunication Union. High efficiency video coding: Recommendation ITU-T H.265. Technical report, ITU-T Series H: Audiovisual and Multimedia Systems, 2013.
- [46] N. Cho. Fast algorithm and implementation of 2-D discrete cosine transform. *IEEE Transactions on Circuits and Systems*, 38:297–305, 1991.
- [47] D. Le Gall. MPEG: A video compression standard for multimedia applications. *Communications of the ACM*, 34(4):46–58, 1991.
- [48] R. K. W. Chan and M. C. Lee. 3D-DCT quantization as a compression technique for video sequences. *Proceedings of the International Conference on Virtual Systems and MultiMedia*, pages 188–196, 1997.
- [49] T. R. Natarajan and N. Ahmed. On interframe transform coding. *IEEE Transactions on Communications*, 25(11):1323–1329, 1977.
- [50] A. Mulla, J. Baviskar, A. Baviskar, and C. Warty. Image compression scheme based on zig-zag 3D-DCT and LDPC coding. In *Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on, pages 2380–2384, 2014.*
- [51] S. Saponara. Real-time and low-power processing of 3D direct/inverse discrete cosine transform for low-complexity video codec. *Journal of Real-Time Image Processing*, 7:43–53, 2012.
- [52] S. Sawant and D. A. Adjeroh. Balanced multiple description coding for 3D DCT video. *IEEE Transactions on Broadcasting*, 57, 2011.
- [53] N. Bozinović and J. Konrad. Motion analysis in 3D DCT domain and its application to video coding. *Signal Processing: Image Communication*, 20:510–528, 2005.

- [54] N. Bozinović and J. Konrad. Scan order and quantization for 3D-DCT coding. In *Proceedings of SPIE Visual Communications and Image Processing*, volume 5150, page 1205, 2003.
- [55] T. H. Lai and L. Guan. Video coding algorithm using 3-D DCT and vector quantization. In *Proceedings of the International Conference on Image Processing*, volume 1, pages I-741, 2002.
- [56] Y.-L. Chan and W.-C. Siu. Variable temporal-length 3-D discrete cosine transform coding. *IEEE Transactions on Image Processing*, 6(5):758–763, 1997.
- [57] M. Servais and G. de Jager. Video compression using the three dimensional discrete cosine transform (3D-DCT). In *Proceedings of the 1997 South African Symposium on Communications and Signal Processing (COMSIG)*, pages 27–32. IEEE, 1997.
- [58] S. Boussakta. Fast algorithm for the 3-D DCT-II. *IEEE Transactions on Signal Processing*, 52:992–1001, 2004.
- [59] Xi Li, A. Dick, C. Shen, A. van den Hengel, and H. Wang. Incremental learning of 3D-DCT compact representations for robust visual tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35:863–881, 2013.
- [60] L. De Lathauwer, B. De Moor, and J. Vandewalle. 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*, 21(4):1324–1342, 2000.
- [61] L. De Lathauwer and B. De Moor. From matrix to tensor: Multilinear algebra and signal processing. In *Institute of Mathematics and Its Applications Conference Series*, volume 67, pages 1–16. Citeseer, 1998.
- [62] D. G. Northcott. *Multilinear algebra*. Cambridge University Press, 2008.
- [63] D. S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas*. Princeton University Press, 2009.
- [64] V. Bhaskaran and K. Konstantinides. *Image and Video Compression Standards*. Kluwer Academic Publishers, 1997.

- [65] Y. Zeng, G. Bi, and A. Rahim Leyman. New polynomial transform algorithm for multidimensional DCT. *IEEE Transactions on Signal Processing*, 48(10):2814–2821, 2000.
- [66] V. K. Madisetti. *The Digital Signal Processing Handbook*. CRC Press, 2009.
- [67] W. H. Chen, C. Smith, and S. Fralick. A fast computational algorithm for the discrete cosine transform. *IEEE Transactions on Communications*, 25:1004–1009, 1977.
- [68] B. G. Lee. A new algorithm for computing the discrete cosine transform. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-32:1243–1245, 1984.
- [69] Z. Wang. Fast algorithms for the discrete W transform and for the discrete Fourier transform. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-32:803–816, 1984.
- [70] H. S. Hou. A fast recursive algorithm for computing the discrete cosine transform. *IEEE Transactions on Acoustic, Signal, and Speech Processing*, 6:1455–1461, 1987.
- [71] C. Loeffler, A. Ligtenberg, and G. S. Moschytz. Practical fast 1-D DCT algorithms with 11 multiplications. *ICASSP International Conference on Acoustics, Speech, and Signal Processing*, 2:988–991, 1989.
- [72] E. Feig and S. Winograd. Fast algorithms for the discrete cosine transform. *IEEE Transactions on Signal Processing*, 40:2174–2193, 1992.
- [73] M. T. Heideman. *Multiplicative Complexity, Convolution, and the DFT*. Signal Processing and Digital Filtering. Springer-Verlag, 1988.
- [74] C.-K. Fong and W.-K. Cham. LLM integer cosine transform and its fast algorithm. *IEEE Transactions on Circuits and Systems for Video Technology*, 22:844–854, 2012.
- [75] Y.-J. Chen, S. Orintara, and T. Nguyen. Video compression using integer DCT. In *Proceedings of International Conference on Image Processing*, volume 2, pages 844–845. IEEE, 2000.

- [76] Y. Yokotani, R. Geiger, G. D. T. Schuller, S. Oraintara, and K. R. Rao. Lossless audio coding using the intMDCT and rounding error shaping. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(6):2201–2211, 2006.
- [77] L. O. Hnativ. Integer cosine transforms: Methods to construct new order 8, 16 fast transforms and their application. *Cybernetics and Systems Analysis*, 50(6):913–929, 2014.
- [78] T. I. Haweel. A new square wave transform based on the DCT. *Signal Processing*, 82:2309–2319, 2001.
- [79] K. Lengwehasatit and A. Ortega. Scalable variable complexity approximate forward DCT. *IEEE Transactions on Circuits and Systems for Video Technology*, 14:1236–1248, 2004.
- [80] R. J. Cintra and F. M. Bayer. A DCT approximation for image compression. *IEEE Signal Processing Letters*, 18:579–582, 2011.
- [81] F. M. Bayer and R. J. Cintra. DCT-like transform for image compression requires 14 additions only. *Electronics Letters*, 48:919–921, 2012.
- [82] S. Bouguezel, M. O. Ahmad, and M. N. S. Swamy. Low-complexity 8×8 transform for image compression. *Electronics Letters*, 44:1249–1250, 2008.
- [83] Z. Wang. Pruning the fast discrete cosine transform. *IEEE Transactions on Communications*, 39:640–643, 1991.
- [84] L. Makkaoui, V. Lecuire, and J. Moureaux. Fast zonal DCT-based image compression for wireless camera sensor networks. In *2nd International Conference on Image Processing Theory Tools and Applications (IPTA)*, pages 126–129. IEEE, 2010.
- [85] V. Lecuire, L. Makkaoui, and J.-M. Moureaux. Fast zonal DCT for energy conservation in wireless image sensor networks. *Electronics Letters*, 48:125–127, 2012.

- [86] N. Kouadria, N. Doghmane, D. Messadeg, and S. Harize. Low complexity DCT for image compression in wireless visual sensor networks. *Electronics Letters*, 49:1531–1532, 2013.
- [87] V. A. Coutinho, R. J. Cintra, F. M. Bayer, S. Kulasekera, and A. Madanayake. A multiplierless pruned DCT-like transformation for image and video compression that requires ten additions only. *Journal of Real-Time Image Processing*, 12(2):247–255, 2016.
- [88] R. J. Cintra, F. M. Bayer, V. A. Coutinho, S. Kulasekera, A. Madanayake, and A. Leite. Energy-efficient 8-point DCT approximations: Theory and hardware architectures. *Circuits, Systems, and Signal Processing*, 35(11):4009–4029, 2016.
- [89] K. Mechouek, N. Kouadria, N. Doghmane, and N. Kaddeche. Low complexity DCT approximation for image compression in wireless image sensor networks. *Journal of Circuits, Systems and Computers*, 2016.
- [90] U. S. Potluri, A. Madanayake, R. J. Cintra, F. M. Bayer, S. Kulasekera, and A. Edirisuriya. Improved 8-point approximate DCT for image and video compression requiring only 14 additions. *IEEE Transactions on Circuits and Systems I: Regular Papers*, PP:1–14, 2014.
- [91] S. Srinivasan. Transformada reversível para compactação de dados bidimensionais com perdas e sem perdas, 2006. BR PI 0506084-2 A.
- [92] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison-Wesley Publishing Company,, 1989.
- [93] A. Martemyanov, N. Terterov, A. Zheludkov, A. Dolgoborodov, V. Semenyuk, E. Neimark, and I. Terterov. Real-time video coding/decoding, 2008. US 7336720 B2.
- [94] S.-Y. Jeong, H.-C. Choi, J. Seo, S.-K. Beack, I.-S. Jang, J.G. Kim, K.-A. Moon, D.-Y. Jang, J.W. Hong, J.-W. Kim, C.-B. Ahn, S.-J. Oh, D.-Gyu Sim, H.-C. Park, and Y.-L. Lee S.-Y. Jeon. Multi-dimensional and multi-stage transforming apparatus and method, 2013. US 8582657 B2.

- [95] T. Uemori, O. Nakagami, M. Yamane, and Y. Yagasaki. Image processing apparatus and image processing method, 2012. US 20120057777 A1.
- [96] A. M. Chiang. Discrete cosine transform processing system, 1992. US 5126962 A.
- [97] M. McCormick, N. Davies, A. Aggoun, and M. C. Forman. Stereoscopic image encoding, 2003. US 6535629 B2.
- [98] L. E. Russo, J. L. Paredes, and G. R. Arce. Bijection mapping for compression/denoising of multi-frame images, 2007. US 7224845 B1.
- [99] S. Watanabe, T. Hayashi, M. Yamane, and T. Uemori. Information processing apparatus and method, image providing system and image providing method, and program, 2012. US 20120251014 A1.
- [100] U. A. Akhatovich, T. A. Konstantinovich, S. N. Nikolaevich, I. S. Mansurovich, I. V. F. Evich, and F. S. Subkhievich. Method of coding and decoding video information based on three-dimensional discrete cosine transformation, 2009. RU 2375838 C1.

REIVINDICAÇÕES

1. “Codificador Tridimensional de Imagem e Vídeo de Baixo Consumo Energético”, caracterizado por ser executado sem operação de ponto flutuante em que a matriz \mathbf{T} em (12) é aplicada às expressões (14) e (16) no bloco de Transformação 3D no esquema de compressão da Figura 3.
2. “Codificador Tridimensional de Imagem e Vídeo de Baixo Consumo Energético”, caracterizado por ser executado sem operação de ponto flutuante que utilize no bloco de Transformação 3D no esquema de compressão da Figura 3 o circuito esquemático rápido exibido na Figura 5, respectivo à matriz \mathbf{T} , para computar as expressões (14) e (16).
3. “Codificador Tridimensional de Imagem e Vídeo de Baixo Consumo Energético”, caracterizado por ser executado sem operação de ponto flutuante em que as matriz \mathbf{T}^\top aplicada às expressões (15) e (25) no bloco de Transformação 3D inversa no esquema de descompressão da Figura 4.
4. “Codificador Tridimensional de Imagem e Vídeo de Baixo Consumo Energético”, caracterizado por ser executado sem operação de ponto flutuante que utilize no bloco de Transformação 3D inversa no esquema de descompressão da Figura 4 o circuito esquemático rápido exibido na Figura 6, respectivo a \mathbf{T}^\top , para computar as expressões (15) e (25).
5. “Codificador Tridimensional de Imagem e Vídeo de Baixo Consumo Energético” de acordo com as Reivindicações 1 e 2, caracterizado por empregar bloco de quantização 3D no esquema da Figura 3 derivado através do método de quantização 3D modificada proposto, sumarizado pelas expressões (21) e (22).
6. “Codificador Tridimensional de Imagem e Vídeo de Baixo Consumo Energético” de acordo com as Reivindicações 3 e 4, caracterizado por empregar bloco de desquantização 3D no esquema da Figura 4 derivado através do método de desquantização 3D modificada proposto, sumarizado pelas expressões (24) e (23).

FIGURAS

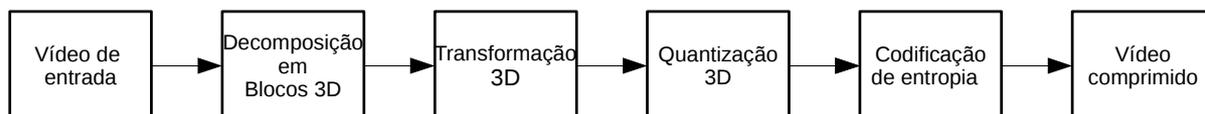


Figura 1

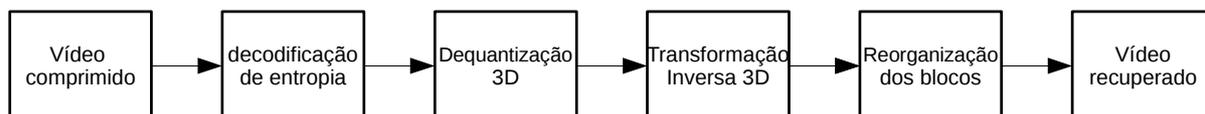


Figura 2

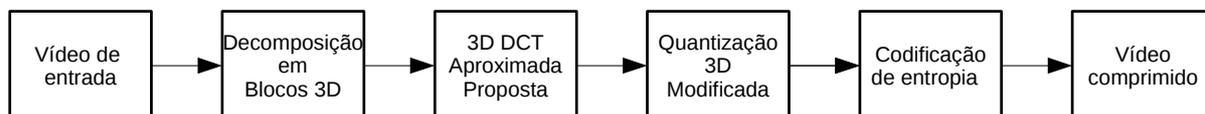


Figura 3

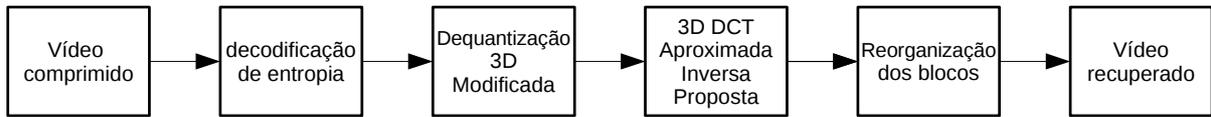


Figura 4

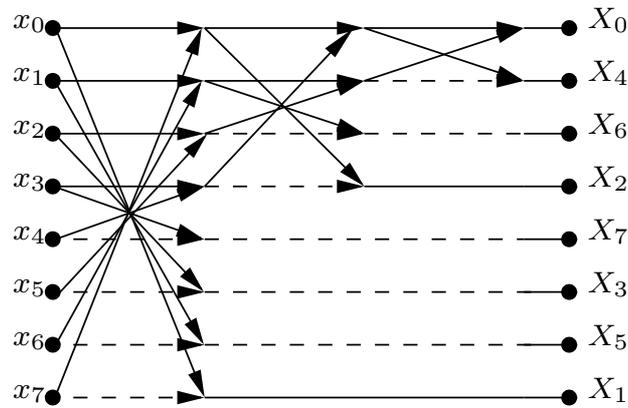


Figura 5

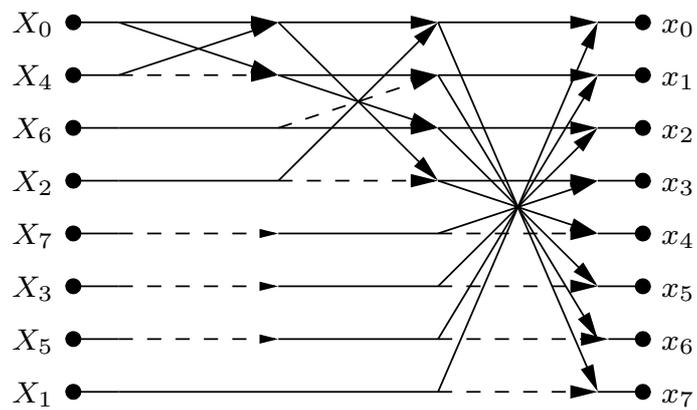


Figura 6

RESUMO

CODIFICADOR TRIDIMENSIONAL DE IMAGEM E VÍDEO DE BAIXO CONSUMO ENERGÉTICO

Esta invenção trata de um esquema de compressão de imagens ou vídeos nos mais variados formatos de dados tridimensionais, apresentando baixo custo computacional e alto desempenho de compactação. O método pode ser aplicado para imagens dinâmicas como vídeo e imagens estáticas tridimensionais. Este esquema se baseia em uma nova transformação tridimensional baseada em uma matriz de transformação associada a uma aproximação para a DCT de baixo custo computacional. A nova transformada tridimensional apresentada evita a necessidade de operações em ponto flutuante durante a etapa de compactação e descompactação dos dados processados. O bloco da transformada usualmente implementada pela transformada discreta do cosseno tridimensional (3D DCT) é substituído pela transformada tridimensional de baixa complexidade apresentada, que possui alta performance de descorrelação de vídeos e outros sinais classificados como processo estacionário Markov do tipo I. Além da substituição do bloco da 3D DCT por uma outra transformada, o bloco de quantização é modificado de modo a tornar a complexidade da transformada apresentada baixa. Além disso, o presente invento apresenta grande eficiência energética por fazer uso de uma matriz com coeficientes apenas em 0 e ± 1 . Isso implica em eliminação da complexidade multiplicativa da etapa de processamento do sinal tridimensional e torna desnecessária a utilização de aritmética de ponto flutuante. Este comportamento da presente dispositivo configura um dispositivo ambientalmente correto - livre tradução de *eco-friendly* ou *green device*.