

## PROJETO DE HARDWARE PARA A TRANSFORMADA INVERSA DCT-II DO PADRÃO VERSATILE VIDEO CODING

BRUNA ROSA GARCIA; BIANCA SILVEIRA; GUILHERME CORRÊA

*Universidade Federal de Pelotas – Video Technology Research Group*  
*{bruna.garcia; bscsilveira; gcorrea}@ufpel.edu.br*

### 1. INTRODUÇÃO

Nos dias atuais, o consumo de conteúdo multimídia tem aumentado significativamente, sendo que a maior parte do tráfego da internet está relacionada ao *streaming* de vídeo (ARULRAJAH, 2021). Segundo CISCO (2018), o conteúdo de vídeo representa cerca de 82% do tráfego global da internet. Com esse aumento da demanda de transmissão de vídeo digital de alta qualidade, a necessidade de desenvolver novas e eficientes técnicas de compressão de vídeo aumentaram também. Nesse cenário, os aplicativos multimídia responsáveis pela codificação e decodificação de vídeos desempenham um papel crucial na obtenção de taxas de compressão mais altas.

O mais recente padrão de codificação de vídeo desenvolvido pela *Joint Video Experts Team* (JVET) é o *Versatile Video Coding* (VVC), finalizado em julho de 2020 (ITU-T e ISO/IEC, 2020), com o objetivo de dobrar a eficiência em relação ao seu predecessor, o *High-Efficiency Video Coding* (HEVC). Neste novo padrão foram introduzidas novas ferramentas distribuídas através dos principais módulos do esquema de codificação tradicional. Entre os novos recursos está a ferramenta *Multiple-Transform Selection* (MTS), que é parte do módulo de transformadas do VVC. Essa ferramenta inclui três tipos de transformadas: a Transformada Discreta do Coseno Tipo II (DCT-II), Tipo VIII (DCT-VIII), e a Transformada Discreta do Seno Tipo VII (DST-VII). As partições dos blocos no VVC podem ser retangulares e quadradas, tendo tamanhos que podem variar de 4x4 até 64x64 para DCT-II e 4x4 até 32x32 para DCT-VIII/DST-VII. Uma outra inovação na ferramenta é que as transformadas podem ser executadas de forma separável, ou seja, pode-se aplicar uma transformada diferente nas direções vertical e horizontal. Por exemplo, a DCT-VIII pode ser aplicada em uma dimensão e a DST-VII na outra dimensão. Além dessas transformadas, o padrão VVC utiliza uma segunda transformada secundária, a *LFNST* (*Low-Frequency Non-Separable Transform*) que pode ser aplicada após a execução da DCT-II primária em blocos intra codificados.

No processo de codificação, o cálculo das transformadas é aplicado com a finalidade de reduzir a redundância espacial entre amostras residuais (diferença entre bloco predito e o bloco de referência) em cada imagem ou quadro de vídeo. A etapa das transformadas é um dos principais módulos do esquema de codificação, sendo responsável pela preparação dos dados a serem modificados pela codificação de entropia, onde de fato acontece a compressão. Apesar de existirem diversas variantes da DCT, a DCT-II é a transformada mais utilizada no processamento digital de imagens e é, atualmente, a base de diversos padrões.

O método inverso da DCT-II é chamado de *Inverse Discrete Cosine Transform Type II* ou Transformada Inversa Discreta do Coseno Tipo II (IDCT-II), e é executada no decodificador. O decodificador é o responsável pela reconstrução

dos dados, os quais são utilizados como referência para o cálculo residual ou estimação de movimento.

Visto que para obter tal eficiência nas taxas de compressão, a complexidade e o consequente custo computacional do codificador e decodificador aumentam, a implementação de hardware dedicados, que otimizam o tempo de codificação e decodificação, torna-se decisivo. Este trabalho, portanto, apresenta um projeto de hardware dedicado para o cálculo da transformada inversa IDCT-II do decodificador do padrão VVC. A arquitetura é implementada apenas para uma dimensão (1D), suportando blocos de tamanhos 4x4 até 32x32.

## 2. METODOLOGIA

Este trabalho propõe uma arquitetura de hardware dedicada para a transformada inversa da DCT-II do padrão de codificação de vídeo VVC. A arquitetura é capaz de processar tamanhos de blocos 4x4 até 32x32 para uma dimensão como pode ser visto na Figura 1.

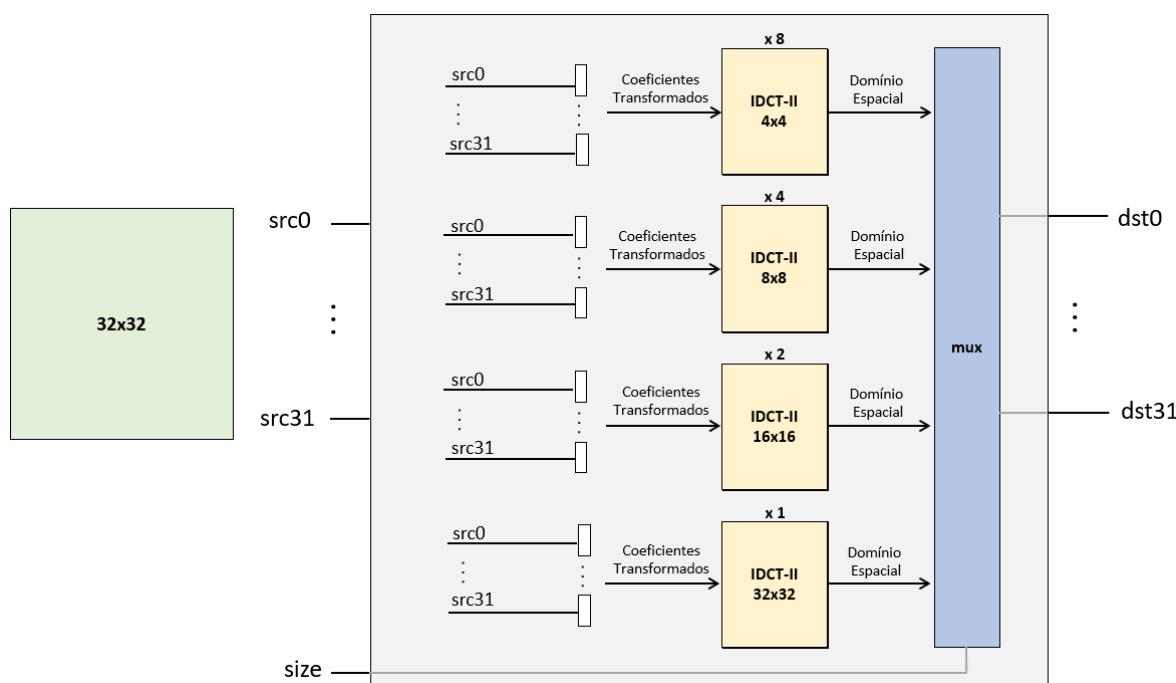


Figura 1 – Arquitetura Proposta para o cálculo da IDCT-II do VVC.

A arquitetura recebe trinta e duas entradas (src0 a src31) que correspondem às amostras do bloco de coeficientes transformados e trinta e duas saídas que corresponderão às colunas resultantes do cálculo (dst0 a dst31). A entrada *SIZE* realiza o controle de seleção do tamanho do bloco escolhido. A arquitetura possui oito módulos de quatro entradas para a realização do cálculo de blocos de tamanho 4x4; quatro módulos de oito entradas para blocos de tamanho 8x8; dois módulos de dezesseis entradas para blocos de tamanho 16x16 e, por fim, um módulo de trinta e duas entradas para blocos de tamanho 32x32. Todas as saídas são conectadas a um multiplexador, cuja saída será aquela correspondente ao tamanho de bloco previamente selecionado pelo software para realização do cálculo (através da entrada *SIZE*).

### 3. RESULTADOS E DISCUSSÃO

A arquitetura foi descrita em linguagem de descrição de hardware HDL (*Hardware Description Language*) utilizando o ambiente ModelSim 6.5b Starter Edition da Altera. A implementação da IDCT-II foi obtida a partir dos cálculos encontrados no software de referência VTM. Para cada tamanho de bloco (4x4, 8x8, 16x16 e 32x32) foi realizada a validação com os dados extraídos do VTM, referentes aos coeficientes de entrada e saída para conferência dos resultados finais.

Para o cálculo das frequências alvo utilizadas na síntese foram consideradas resoluções Full HD (1920x1080 pixels), UHD 4K (3840x2160 pixels) e 8K (7680x4320 pixels), bem como as taxas de amostragem de 30 e 60 quadros por segundo (qps ou fps). O formato de subamostragem considerado para os cálculos foi de 4:2:0, em que para cada quatro componentes de luminância, tem-se uma componente de crominância azul e uma de crominância vermelha no espaço de cores YCbCr.

O cálculo das frequências alvo é baseado na mínima taxa de processamento para diferentes resoluções de vídeo a serem codificados ou decodificados em tempo real (Equação 1).

$$f_{\text{target}} = \frac{\text{resolução} \times \text{subamostragem de cores} \times \text{nº ciclos} \times \text{fps}}{\text{nº de amostras}} \quad (1)$$

A arquitetura foi sintetizada, com dados aleatórios, com a ferramenta Cadence *RTL Compiler* 14.2 com o padrão TSMC e tecnologia 40nm.

Na Tabela I é possível comparar os diferentes resultados referentes a área e dissipação de potência total para cada taxa de processamento e suas respectivas frequências alvo.

**Tabela I – Resultados de Síntese para Diferentes Frequências Alvo.**

Resolução	Frequência (MHz)	Área (kGates)	Potência Total (mW)
Full HD@30 fps	2,916	106,85	3,636
Full HD@60 fps	5,832		7,084
UHD 4K@30 fps	11,664		12,279
UHD 4K@60 fps	23,328		24,397
UHD 8K@30 fps	46,656		44,759
UHD 8K@60 fps	93,312		77,292

Como esperado, a dissipação de potência aumenta conforme a frequência aumenta. Logo, a maior dissipação total é para resoluções UHD 8K com taxa de processamento de 60 fps. Através da síntese também foi verificada a frequência de operação máxima suportada pela arquitetura proposta, que é de 833,33 MHz.

## 4. CONCLUSÕES

A arquitetura desenvolvida propõe um projeto de hardware dedicado para o cálculo da transformada inversa IDCT-II do decodificador do padrão VVC. Ela suporta blocos de tamanho 4x4 até 32x32 e a possibilidade de seleção de tamanho do bloco a ser decodificado, realizando a operação em uma única dimensão. A arquitetura proposta é capaz de processar vídeos de resolução UHD 8K a uma taxa de 60 fps com uma dissipação de potência de 77,29 mW.

Em trabalhos futuros será feita a adaptação da arquitetura apresentada para o processamento de duas dimensões e, além disso, o desenvolvimento das demais transformadas incluídas na ferramenta MTS do software de referência.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

ARULRAJAH, Naren. **The World is in a State of Flux. Is Your Website Traffic, Too?** Oral The Help, 26 fev. 2021. Acessado em 15 ago. 2022. Online. Disponível em: <https://www.oralhealthgroup.com/features/the-world-is-in-a-state-of-flux-is-your-website-traffic-too/>

**Cisco Visual Networking Index: Forecast and Trends, 2017 – 2022.** CISCO, nov. 2018. Acessado em 15 ago. 2022. Online. Disponível em: [www.cisco.com](http://www.cisco.com)

ITU-T., ISO/IEC. **Versatile Video Coding.** ITU-T Recommendations, 29 ago. 2020. Disponível em: <https://handle.itu.int/11.1002/1000/14336>. Acesso em: 15 ago. 2022.

XIN, Zhao *et al.* Transform Coding in the VVC Standard. **IEEE Transactions on Circuits and Systems for Video Technology**, [s. l.], ano 2021, v. 31, ed. 10, p. 3878-3890, 2021. DOI 10.1109/TCSVT.2021.3087706.