

PROJETO DE HARDWARE PARA O MÓDULO DA MTS DO PADRÃO DE CODIFICAÇÃO DE VÍDEO VVC

BIANCA SILVEIRA¹; DANIEL PALOMINO²; CLÁUDIO DINIZ³;
GUILHERME CORREA⁴

¹Universidade Federal de Pelotas, ViTech – bscsilveira@inf.ufpel.edu.br

²Universidade Federal de Pelotas, ViTech - dpalomino@inf.ufpel.edu.br

³Universidade Federal do Rio Grande do Sul, PGMICRO – cmdiniz@inf.ufrgs.br

⁴Universidade Federal de Pelotas, ViTech – gcorrea@inf.ufpel.edu.br

1. INTRODUÇÃO

A crescente popularidade dos vídeos digitais tem sido acompanhada por um notável aumento na demanda. Nos últimos anos, o ambiente de multimídia passou por uma significativa transformação digital, com uma rápida disseminação de aplicativos voltados para a manipulação de vídeos digitais. Entretanto, é importante ressaltar que um vídeo com duração de 5 minutos, exibindo resolução Full High Definition (FHD) de 1920x1080 pixels, operando a uma taxa de 60 quadros por segundo (60fps) e com uma profundidade de cor de 24 bits (3 bytes) por pixel, requer aproximadamente 112 gigabytes (GB) de espaço de armazenamento. Portanto, o processo de compressão de vídeo assume uma importância extrema para viabilizar eficientemente a transmissão desses conteúdos.

O Versatile Video Coding (VVC) (ITU-T and ISO/IEC, 2020) representa um dos mais recentes padrões de compressão de vídeo, cujo desenvolvimento foi concluído em julho de 2020. Dentre suas várias inovações, destaca-se a introdução da ferramenta *Multiple Transform Selection* (MTS), que engloba a Transformada Discreta do Seno tipo VII (DST-VII), a Transformada Discreta do Cosseno tipo VIII (DCT-VIII) e a Transformada Discreta do Cosseno tipo II (DCT-II). Além disso, os blocos residuais atingem dimensões de até 64x64 para a DCT-II e até 32x32 para a DCT-VIII e DST-VII. Através da incorporação de transformações separáveis, essas transformadas podem ser aplicadas de forma independente nas direções horizontal e vertical. Por exemplo, quando se opta pela DST-VII na orientação vertical, tanto a DST-VII quanto a DCT-VIII são viáveis na orientação horizontal, e vice-versa.

A incorporação da MTS acarreta um incremento na complexidade do módulo de Transformadas do VVC, tornando crucial a minimização da complexidade das operações e a implementação de hardware especializado para viabilizar a compressão em tempo real. O suporte à MTS para múltiplas transformadas pode impactar a alocação de memória e a utilização de recursos lógicos, tornando fundamental a concepção de um projeto de alto desempenho que esteja em conformidade com as restrições de hardware, conforme discutido em KAMMOUN (2019).

2. METODOLOGIA

No codificador VVC, a ferramenta MTS introduziu uma capacidade inovadora de utilizar múltiplos tipos de transformadas em um único bloco de resíduos. Isso difere do seu antecessor, o HEVC, que apenas usava a DCT-II. No VVC, as transformadas DCT-VIII e DST-VII foram adicionadas. Além disso, o VVC oferece uma flexibilidade significativa nos tamanhos dos blocos transformados, que

podem ser quadrados ou retangulares. Esses tamanhos variam de 4x4 a 64x64 para a DCT-II e de 4x4 até 32x32 para as transformadas DCT-VIII e DST-VII. Quando se trata de blocos retangulares, os tamanhos podem ser configurados como $N/2 \times N$, $N \times N/2$, $N/4 \times N$ ou $N \times N/4$.

A MTS envolve a aplicação de transformadas unidimensionais (1D) nas direções horizontal e vertical em um bloco de dimensões $N \times M$, resultando na obtenção de um bloco final também de dimensões $N \times M$, porém transformado. O elemento característico da MTS está na possibilidade de empregar transformadas distintas em cada uma das orientações, sendo que tais escolhas são determinadas pelo codificador. O controle da funcionalidade da MTS é efetuado mediante a ativação ou desativação de flags específicas, as quais são encontradas nos parâmetros de sequência de vídeo (Sequence Parameter Set - SPS) do software de referência VTM 10.0 (FRAUNHOFER).

Neste estudo, foram desenvolvidas implementações de arquiteturas unidimensionais (1D) para os três tipos de transformadas permitidas na MTS (DCT-II, DCT-VIII e DST-VII). Adicionalmente, foi elaborada uma arquitetura unificada para as diversas combinações possíveis de transformadas bidimensionais (2D). As arquiteturas propostas suportam uma variedade de dimensões de blocos, incluindo 4x4, 4x8, 4x16, 4x32, 8x4, 8x8, 8x16, 8x32, 16x4, 16x8, 16x16, 16x32, 32x4, 32x8, 32x16 e 32x32.

Com base na análise realizada no software de referência VTM, é possível observar que os coeficientes das matrizes menores da DCT-II, estão presentes nas matrizes maiores, assim algumas operações podem ser reutilizadas das matrizes menores nas matrizes maiores.

A técnica de reutilização de coeficientes empregada na DCT-II não pode ser aplicada às transformadas DCT-VIII e DST-VII devido à ausência de quaisquer semelhanças compartilhadas entre os coeficientes em diferentes tamanhos de matriz. Conseqüentemente, uma abordagem alternativa deve ser desenvolvida para essas transformações específicas. No entanto, uma análise mais detalhada revela que os coeficientes de DCT-VIII e DST-VII possuem uma peculiar similaridade: eles frequentemente têm valores espelhados com sinais invertidos, portanto, foi desenvolvida uma única arquitetura que permite a reutilização de operadores na implementação destas duas transformadas.

A arquitetura proposta tem como objetivo processar um bloco residual de tamanho 32x32, que pode ser subdividido em diferentes subblocos quadrados e retangulares, de acordo com as especificações permitidas pelo VVC. O bloco de 64x64 não foi considerado nesta arquitetura, uma vez que a DCT-II é sempre escolhida para essa opção de tamanho de bloco. A Figura 1 ilustra a arquitetura de múltiplas transformadas proposta.

Para realizar o cálculo 1D para o bloco residual completo de 32x32, a arquitetura opera ao longo de um total de 36 ciclos de clock. Esse período inclui 32 ciclos para a realização dos cálculos em si e uma latência inicial de 4 ciclos necessária para preencher o pipeline. A partir do quarto ciclo de clock, o bloco intermediário de 32x32 fornece valores acessíveis para as quatro primeiras linhas. Conseqüentemente, essas saídas se tornam imediatamente disponíveis para serem utilizadas como entradas para realizar os cálculos necessários para a segunda transformada 1D. A arquitetura apresenta uma latência inicial de 68 ciclos de clock para realizar as duas dimensões do bloco de 32x32.

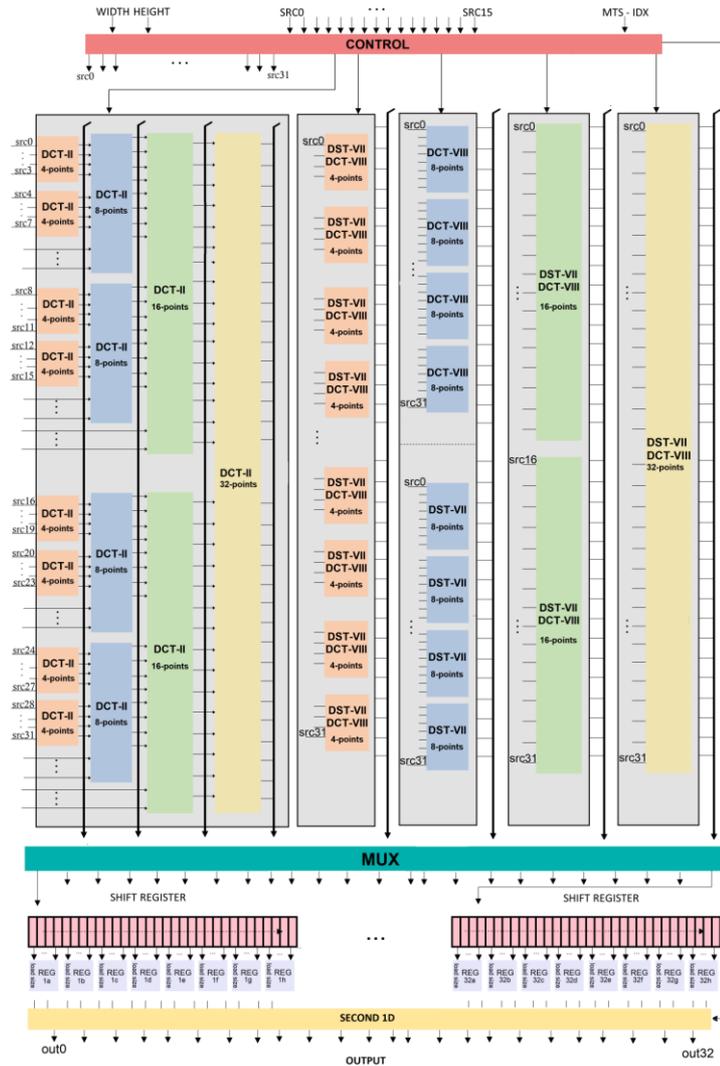


Figura1. Arquitetura Proposta

3. RESULTADOS E DISCUSSÃO

A arquitetura proposta foi descrita utilizando a linguagem VHDL e foi posteriormente sintetizada para ASIC, fazendo uso da ferramenta de software Cadence RTL Compiler 14.02. A síntese foi realizada considerando a tecnologia de células-padrão da TSMC de 40 nm. Entradas de teste geradas aleatoriamente pela ferramenta de síntese foram empregadas durante o processo de síntese.

A fim de garantir que a taxa de processamento necessária para cada resolução de vídeo fosse atingida, foram calculadas as frequências alvo para cada resolução utilizada. O escopo deste estudo abrangeu as seguintes resoluções de vídeo e taxas de quadros (frames por segundo - fps): HD 720p a 30 fps e 60 fps, FHD 1080p a 30 fps e 60 fps, UHD 4K a 30 fps e 60 fps, e 8K a 30 fps.

TABELA 2. Resultados de Síntese

Resolution	Frequency (MHz)	Area (kgates)	Power (mW)
HD@30fps	13.82	2593.53	244.23
HD@60fps	27.65		450.00
FHD@30fps	31.10		569.99
FHD@60fps	62.21		827.20
4K@30fps	124.41		1477.34
4K@60fps	248.83		1988.99
8K@30fps	497.66		3226.50

A Tabela 2 apresenta resultados referentes à área do circuito e à potência total dissipada, conforme relatado na síntese, para cada taxa de transferência e sua frequência correspondente. Observa-se na tabela, conforme esperado, um aumento na dissipação total de potência à medida que a frequência aumenta. A maior dissipação de potência é observada na taxa de transferência de 8K@30fps.

O design proposto aborda de forma abrangente a ferramenta MTS incluída no codec VVC, oferecendo cobertura completa para blocos residuais de até 32x32 amostras ou menores (quadrados e retangulares). Além disso, ele oferece suporte extensivo para todas as combinações possíveis de transformadas 2D permitidas pela ferramenta MTS, possibilitando o processamento de vídeos em tempo real com resolução de até 8K a 30 fps.

4. CONCLUSÕES

Neste estudo, é apresentado um novo projeto de hardware de múltiplas transformadas destinado ao módulo de Transformadas do VVC. A arquitetura proposta permite a exploração de todas as combinações de transformadas da MTS para qualquer tipo de bloco residual, enquanto também oferece suporte para tamanhos de bloco de $M \times N$ (onde $M, N = 4, 8, 16, 32$). Com base na revisão da literatura realizada, podemos afirmar que este é o primeiro projeto publicado para o módulo de Transformadas do VVC que abrange todas as combinações permitidas pela MTS, juntamente com as dimensões de bloco mencionadas anteriormente. Com uma dissipação de potência de 3226,50 mW, a arquitetura tem a capacidade de processar vídeos em tempo real com resolução de até 8K@30fps. Em futuras pesquisas, a intenção é investigar estratégias para otimizar ainda mais a área e a dissipação de potência do módulo de Transformadas.

5. REFERÊNCIAS BIBLIOGRÁFICAS

ITU-T and ISO/IEC, “Versatile Video Coding,” ITU-T Recommendation H.266 and ISO/IEC 23090-3, 2020.

Kammoun, A.; Hamidouche, W.; Philipp, P.;Belghith, F.; Massmoudi, N.; and Nezan, J. F.. “Hardware acceleration of approximate transform module for the versatile video coding standard,” in 2019 27th European Signal Processing Conference (EUSIPCO). IEEE, 2019, pp. 1–5.

FRAUNHOFER. VTM reference software, VVC test model (VTM). Disponível em: <https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM>.