

DESIGN DE HARDWARE DE ALTA PERFORMANCE PARA A PREDIÇÃO ANGULAR INTRA QUADRO DO VVC

VINICIUS BORGES¹; MURILO PERLEBERG¹;
MARCELO PORTO¹; LUCIANO AGOSTINI¹

¹*Universidade Federal de Pelotas – {vdaborges, mrperleberg, porto, agostini}@inf.ufpel.edu.br*

1. INTRODUÇÃO

A sociedade atual depende de muitos dispositivos que fornecem serviços de vídeo digital principalmente para fins de trabalho, educação e entretenimento. Essa alta demanda cria a necessidade de uma nova geração de padrões de codificação de vídeo, com melhorias na eficiência de codificação quando comparados aos antecessores.

O padrão de codificação de vídeo de última geração é o Versatile Video Coding (VVC) (SALDANHA et al., 2020). Foi concluído pela ISO e ITU-T em julho de 2020. Uma das principais melhorias do VVC está relacionada ao particionamento de blocos. O tamanho máximo da Coding Tree Unit (CTU) foi aumentado para 128x128. O particionamento dos tamanhos de bloco foi alterado com o uso do Quadtree com estrutura de Multi-type Tree (QTMT) aninhada, que permite um número maior de tamanhos de Coding Unit (CU) e um esquema de particionamento mais flexível (HUANG et al., 2020). O VVC também tem várias melhorias nas ferramentas de codificação, mas aumentando expressivamente o custo computacional do codificador e decodificador.

Este trabalho é focado na ferramenta AIP. A ferramenta VVC AIP usa um conjunto de amostras de referência de CUs previamente processadas para realizar a predição da CU que está sendo codificada. Essas amostras pertencem à CU vizinha acima e à esquerda da CU que está sendo predita. A ferramenta AIP suporta um total de 17 tamanhos de CU e 67 modos angulares.

Desde o lançamento do VVC, apenas dois trabalhos foram publicados na literatura propondo hardware dedicado para a ferramenta de predição intra AIP deste padrão. O trabalho em (AZGIN; KALALI; HAMZAOGLU, 2019) apresenta um hardware dedicado para predição intra VVC, apresentando uma implementação FPGA Virtex 7 para modos VVC AIP suportando tamanhos de CU 4x4, 8x8, 16x16 e 32x32. A solução em (BORGES et al., 2023) apresenta uma heurística amigável ao hardware proposta para o AIP, que ignora os modos angulares menos usados e alguns dos tamanhos de CU suportados, visando reduzir o esforço computacional do AIP, além de apresentar um projeto arquitetônico para o VVC AIP usando a heurística proposta.

O principal objetivo deste trabalho foi projetar um hardware AIP completo e de alto rendimento visando o codificador VVC. A arquitetura projetada é capaz de atingir a máxima eficiência de codificação possível no AIP, já que nenhuma das avaliações é pulada, o que é feito em outros trabalhos como uma forma de reduzir o esforço computacional do AIP. De fato, mesmo o software de referência VVC Test Model (VTM) (BOSSEN et al., 2019) usa algumas heurísticas para evitar a avaliação de todos os modos angulares do AIP (PFAFF et al., 2021). A arquitetura apresentada também foi projetada visando vídeos de alta resolução, então um alto rendimento é necessário. Para atingir esse objetivo, todos os 67 modos angulares são processados em paralelo e dentro de cada unidade de modo angular, uma linha do CU é processada em paralelo, independentemente do tamanho do CU.

2. METODOLOGIA

O hardware apresentado neste trabalho foi dividido em 67 unidades de processamento: unidade planar, unidade DC e 65 unidades direcionais, como pode ser visto na Figura 1. Cada unidade realiza o processamento da CU prevista para um AIP diferente. Como este trabalho está focado em alto rendimento, todas as 67 unidades trabalham em paralelo. A unidade planar realiza cálculos planares, a unidade DC é responsável pelos cálculos do modo DC e cada uma das 65 unidades direcionais é responsável por processar seu respectivo modo angular, totalizando os 67 modos AIP. As 67 unidades de processamento foram projetadas para prever uma linha da CU atual em cada ciclo de clock, independentemente do tamanho da CU. Isso significa que a arquitetura pode processar uma linha de uma CU 64x64 ou uma linha de uma CU 4x4 em cada ciclo de clock.

Basicamente, cada unidade recebe um sinal de controle para sinalizar que a operação deve começar e um conjunto de amostras de referência das CUs vizinhas esquerda e superior. A arquitetura foi projetada para suportar 8 bits por amostra, representado na Figura 1. Planar e DC têm entradas diferentes em comparação com unidades direcionais porque realizam cálculos diferentes. Finalmente, após cada unidade gerar o bloco previsto, o bloco residual de cada modo é obtido através de subtratores, como pode ser visto na Figura 1.

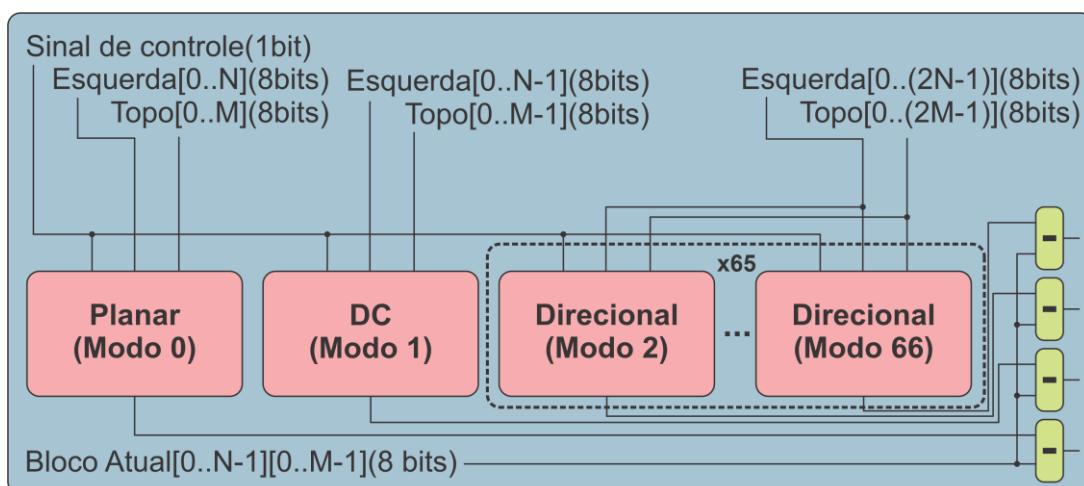


Figura 1. Parte topo da arquitetura proposta para os modos angulares da predição intra quadro do VVC.

A unidade de processamento planar prediz a CU utilizando alguns pesos baseado nas amostras vizinhas, esses pesos serão definidos de acordo com a proximidade entre a amostra a ser predita e a amostra vizinha. Utilizando as amostras com seus respectivos pesos, é então feita uma média ponderada entre as amostras vizinhas utilizadas para gerar a amostra predita.

A unidade de processamento DC faz uma soma total de todas as amostras vizinhas disponíveis gerando um valor total das amostras. Esse valor total é dividido pela quantidade de amostras vizinhas utilizadas, gerando a média aritmética dessas amostras que pode ser chamado de valor DC. Esse valor DC é distribuído para todas as amostras da CU a ser predita.

As unidades de processamento direcionais utilizam certas amostras vizinhas para predizer a CU baseado em um ângulo definido pelo VVC. Devido ao seu ângulo, as unidades direcionais 2-34 utilizam apenas amostras vizinhas à esquerda da CU, enquanto que as unidades direcionais 35-66 utilizam apenas amostras acima da CU.

3. RESULTADOS E DISCUSSÃO

A arquitetura projetada foi descrita em VHDL e sintetizada visando a tecnologia TSMC 40nm usando a ferramenta de compilação Cadence RTL. Os resultados de potência consideraram a atividade de chaveamento padrão definida pela ferramenta de síntese.

A Tabela I resume os resultados da síntese (última coluna) e também apresenta os resultados de alguns trabalhos relacionados. A frequência máxima de operação atingida foi de 250 MHz. A frequência de operação para processar vídeos HD1080p@30fps é de 131,3 MHz, conforme discutido anteriormente.

Além dos resultados da síntese, a Tabela I também apresenta resultados de eficiência de codificação em termos de taxa BD (BJONTEGAARD, 2008). Esta métrica é usada para comparar diferentes implementações de um codificador de vídeo. Ela considera, para a mesma qualidade objetiva, quais são os impactos no número de bits necessários para representar o vídeo. Valores positivos são perdas, pois mais bits serão necessários para representar o mesmo vídeo. Os resultados na Tabela I consideraram o VTM como âncora. Como nossa arquitetura suporta mais modos do que o VTM, de fato nossa taxa de BD será negativa, apresentando alguns ganhos, mas ainda não temos esses resultados experimentais.

Uma comparação justa com os trabalhos relacionados é difícil, pois esses trabalhos focam em tecnologias e experimentos distintos para avaliar a eficiência da codificação. Além disso, todos os trabalhos relacionados apresentam alguma perda na eficiência da codificação, pois usam heurísticas para reduzir o custo computacional do AIP.

Apenas os trabalhos (AZGIN; KALALI; HAMZAOGLU, 2019) e (BORGES et al., 2023) visaram a predição intra VVC, e o trabalho (AZGIN; KALALI; HAMZAOGLU, 2019) apresenta apenas resultados FPGA e não apresenta potência nem eficiência de codificação, evitando qualquer comparação relevante.

O trabalho (BORGES et al., 2023) é nosso trabalho anterior e, embora se concentre no VVC, essa arquitetura suporta apenas 18 modos AIP e cinco tamanhos de CU quadráticos, o que implica em uma perda de taxa de BD de 8,62%, o que é uma perda importante em termos de eficiência de codificação. Conforme discutido anteriormente, a solução apresentada neste trabalho não gera nenhuma degradação em termos de taxa de BD, uma vez que todos os tamanhos de CU (17) e modos AIP (67) são suportados. Naturalmente, esse aumento nos modos suportados e tamanhos de CU impacta diretamente nos resultados de hardware alcançados. Como se pode notar na Tabela I, os resultados de área e potência são muito menores em (BORGES et al., 2023). Mas considerando que a arquitetura atual tem 3,7 vezes mais unidades de processamento, e cada unidade de processamento pode suportar 3,4 vezes mais tamanhos de CU, o aumento de área e potência foi menor do que o esperado.

Tabela I: Resultados de Síntese e Comparação.

| | Azgin 2019 | Borges [7] | Este Trabalho |
|------------------------------------|--------------------|-------------------|----------------------|
| Padrão | VVC | VVC | VVC |
| Modos | 65 | 18 | 67 |
| Tamanhos de CU | 4 | 5 | 17 |
| BD-rate | - | 8,62% | < 0% |
| Tecnologia | FPGA Virtex 7 28nm | TSMC 40nm | TSMC 40nm |
| Área (gate-equivalente) | - | 1.453 k | 13.508,4 k |
| Frequência para 1080p@30fps | 119 MHz | 75,8 MHz | 131,3 MHz |
| Potência para 1080p@30fps | - | 91,65 mW | 755,18 mW |

4. CONCLUSÕES

Este trabalho apresentou um projeto de hardware completo e de alto rendimento para a ferramenta de predição angular intra-frame VVC. Esta arquitetura dedicada suporta todos os 67 modos angulares e todos os 17 tamanhos de CU suportados pelo padrão VVC para a ferramenta AIP e pode processar vídeos 1080p em tempo real. Este projeto de hardware pode atingir a máxima eficiência de codificação possível para esta ferramenta de codificação. Para atingir o alto rendimento, 67 unidades de hardware paralelas e dedicadas foram projetadas, cada uma especializada em um dos modos angulares definidos na ferramenta AIP. Dentro de cada unidade de processamento, o paralelismo aplicado permite o processamento de uma linha de CU a cada ciclo de clock, independentemente do tamanho da CU. Os resultados da síntese mostraram que, para uma frequência de operação de 131,3 MHz, uma área de 13.508,4k gates e uma dissipação de potência de 755,18 mW. Nossa solução é a primeira na literatura que não causa nenhuma degradação da eficiência de codificação.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- AZGIN, H.; KALALI, E.; HAMZAOGLU, I. **An Efficient FPGA Implementation of Versatile Video Coding Intra Prediction**. 2019 22nd Euromicro Conference on Digital System Design (DSD). *Anais...* Em: 2019 22ND EUROMICRO CONFERENCE ON DIGITAL SYSTEM DESIGN (DSD). ago. 2019. Disponível em: <<https://ieeexplore.ieee.org/document/8875183>>. Acesso em: 12 fev. 2024
- BJONTEGAARD, G. Improvements of the BD-PSNR model. jul. 2008.
- BORGES, V. et al. **Efficient Architecture for VVC Angular Intra Prediction based on a Hardware-Friendly Heuristic**. 2023 IEEE 14th Latin America Symposium on Circuits and Systems (LASCAS). *Anais...* Em: 2023 IEEE 14TH LATIN AMERICA SYMPOSIUM ON CIRCUITS AND SYSTEMS (LASCAS). fev. 2023. Disponível em: <<https://ieeexplore.ieee.org/document/10108393>>. Acesso em: 12 fev. 2024
- BOSSEN, F. et al. JVET common test conditions and software reference configurations for SDR video. **Joint Video Experts Team (JVET) of ITU-T SG**, v. 16, p. 19–27, 2019.
- HUANG, Y.-W. et al. A VVC Proposal With Quaternary Tree Plus Binary-Ternary Tree Coding Block Structure and Advanced Coding Techniques. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 30, n. 5, p. 1311–1325, maio 2020.
- PFAFF, J. et al. Intra Prediction and Mode Coding in VVC. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 31, n. 10, p. 3834–3847, out. 2021.
- SALDANHA, M. et al. **An Overview of Dedicated Hardware Designs for State-of-the-Art AV1 and H.266/VVC Video Codecs**. 23 nov. 2020.