

ARQUITETURA DE BAIXO CUSTO EM ÁREA PARA A PREDIÇÃO INTRAQUADRO MULTIPADRÃO H.264/AVC E HEVC COM FOCO EM VÍDEOS DE ALTA RESOLUÇÃO

**MARCEL CORRÉA^{1,2}; MARCELO PORTO²; BRUNO ZATT²;
LUCIANO AGOSTINI²**

¹*Instituto Federal Sul-rio-grandense – marcelcorrea@ifsul.edu.br*

²*Universidade Federal de Pelotas – Grupo de Arquiteturas e Circuitos Integrados
{porto; zatt; agostini}@inf.ufpel.edu.br*

1. INTRODUÇÃO

A codificação de vídeo é uma área de grande interesse industrial e acadêmico atualmente devido ao crescente aumento do número de aplicações e dispositivos eletrônicos capazes de manipular vídeos digitais de alta resolução. Padrões de codificação de vídeo já estão há vários anos presentes no nosso dia a dia e a não utilização destes para manipular vídeos digitais se tornou impraticável. O motivo disso é que um vídeo digital sem qualquer tipo de compressão necessita de uma quantidade imensa de informação digital para ser representado. Por exemplo, caso nenhuma técnica de codificação de vídeo seja utilizada, um vídeo com resolução de 1920×1080 pixels, duração de 10 minutos, taxa de atualização de 30 quadros por segundo e 24 bits por pixel necessita de aproximadamente 105 GB de informação digital para ser representado e uma banda de 1,5 Gbps para ser transmitido em tempo real.

Um padrão de codificação de vídeo define através de um conjunto de técnicas, basicamente dois procedimentos: uma representação codificada – ou sintaxe – que descreve uma informação visual de maneira comprimida, e um método para decodificar esta sintaxe e reconstruir a informação visual (RICHARDSON, 2003). Estas ferramentas se baseiam em reduzir três tipos de redundâncias presentes nos vídeos digitais, sendo estas: (a) espacial ou intraquadro, caracterizada por semelhanças em regiões vizinhas dentro de um quadro; (b) temporal ou interquadros, presente em regiões similares entre quadros vizinhos; e (c) entrópica, presente na representação binária da informação (RICHARDSON, 2003).

A predição intraquadro – técnica para redução da redundância espacial no próprio domínio espacial –, foi uma inovação do padrão H.264/AVC também adotada pelo padrão estado da arte HEVC. Neste último, o módulo de predição intra pode ser visto como três diferentes etapas: pré-processamento das amostras de referência, predição de amostras e pós-processamento das amostras preditas. Superior ao módulo de predição intra do padrão antecessor H.264 que permite apenas 9 modos de predição sobre blocos de tamanho 4×4 ou 16×16 (RICHARDSON, 2003), o padrão HEVC suporta um total de 35 modos de predição sobre tamanhos de bloco que variam de 4×4 até 32×32. Destes modos, 33 são utilizados para a predição de blocos em regiões com características direcionais, e 2 são utilizados para regiões homogêneas.

Este trabalho propõe uma solução arquitetural para o módulo de predição intraquadro multipadrão H.264 e HEVC. Esta arquitetura foi projetada para ser capaz de processar blocos de tamanho 4×4 com os modos Planar, DC, Horizontal e Vertical com baixo grau de paralelismo para redução de custo em área, embora mantendo uma taxa de processamento adequada para a codificação de vídeos de alta resolução.

2. METODOLOGIA

Para o desenvolvimento arquitetural, os padrões e seus algoritmos foram estudados com o objetivo de simplificá-los para uma descrição em hardware. Após esta etapa, as duas partes da arquitetura foram descritas na linguagem VHDL. Estas duas partes são: (a) módulo de predição e filtragem, e (b) módulo de comparação de candidatos. A arquitetura completa pode ser vista na Figura 1, onde os dois módulos citados anteriormente estão destacados pelos retângulos tracejados.

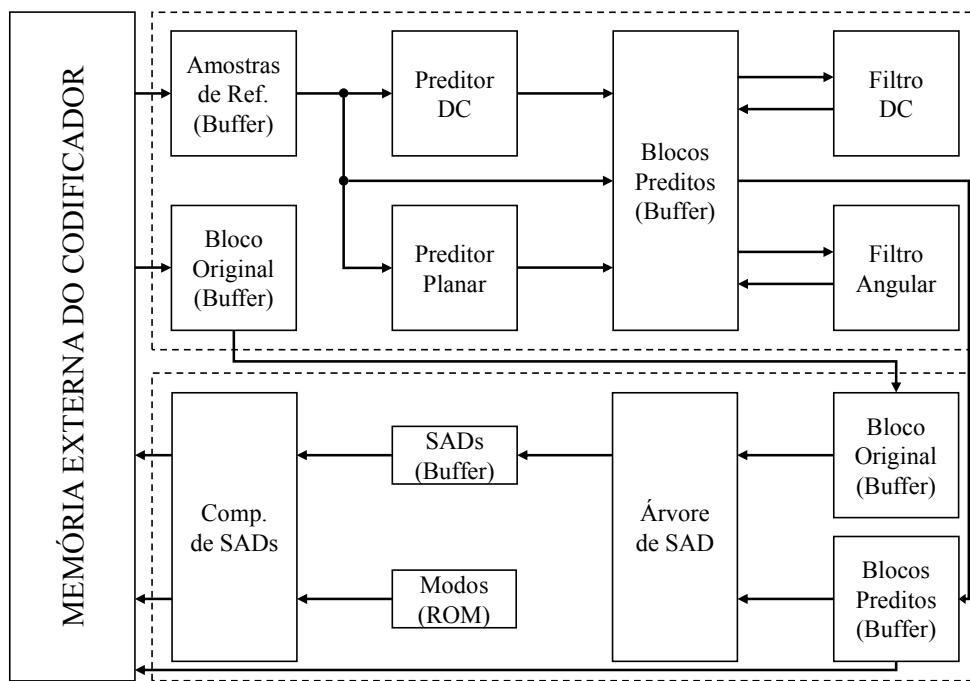


Figura 1 – Arquitetura Completa.

A predição de um bloco 4×4 necessita 11 amostras de referências. Considera-se que estas estão organizadas em um array 2D, sendo $p[0..4][1]$ as cinco amostras horizontais, $p[-1][0..4]$ as cinco verticais e $p[-1][-1]$ a diagonal.

Cada amostra na predição Planar é obtida através de duas interpolações lineares (horizontal e vertical) e da média simples destas interpolações, conforme mostra as equações (1-3) respectivamente. A arquitetura capaz de obter uma amostra predita Planar pode ser vista na Figura 2.

$$p_H[x][y] = (3 - x) * p[-1][y] + (x + 1) * p[4][-1] \quad (1)$$

$$p_V[x][y] = (3 - y) * p[x][-1] + (y + 1) * p[-1][4] \quad (2)$$

$$p[x][y] = (p_H[x][y] + p_V[x][y] + 4) \gg 3 \quad (3)$$

Na predição DC, o bloco obtido é totalmente homogêneo e portanto todas amostras preditas são idênticas. Esta amostra é chamada dcVal e é obtida através da média simples entre todas amostras de referência horizontais e verticais, conforme mostra a equação (4). A arquitetura capaz de obter a amostra dcVal pode ser vista na Figura 3.

$$dcVal = \left(4 + \sum_{i=0}^3 \sum_{j=0}^3 p[i][j] \right) \gg 3 \quad (4)$$

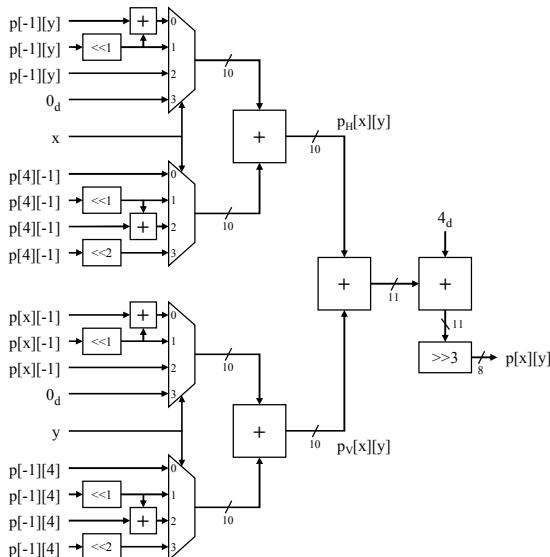


Figura 2 – Predictor Planar.

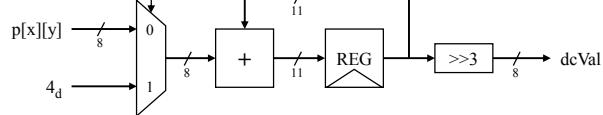


Figura 3 – Predictor DC.

Para a etapa de filtragem, considera-se que o bloco predito está organizado no mesmo array 2D das amostras de referência, mas nos índices $p[0...3][0...3]$.

O Filtro DC deve substituir as bordas lateral esquerda e superior do bloco predito DC para eliminar as descontinuidades geradas. Este tratamento é feito através do filtro de três taps mostrado na Figura 4, onde as entradas ref_a e ref_b serão $p[-1][0]$ e $p[0][-1]$ para a filtragem da amostra $p[0][0]$, e serão dcVal e ref_adj para as demais.

O Filtro Angular deve substituir a borda superior do bloco predito pelo modo horizontal e a borda lateral esquerda do bloco predito pelo modo vertical. Este tratamento é feito pelo filtro mostrado na Figura 5.

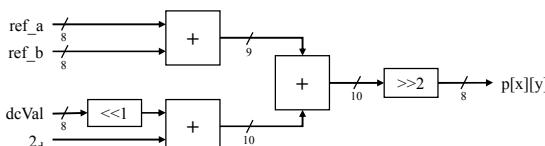


Figura 4 – Filtro DC.

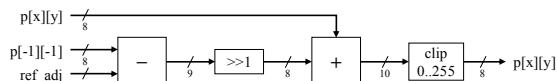


Figura 5 – Filtro Angular.

Após o cálculo dos quatro blocos preditos, o módulo de comparação de candidatos deve escolher qual será entregue ao codificador. Neste trabalho, foi adotada o cálculo da métrica SAD (*Sum of Absolute Differences*), que é a soma do valor absoluto das diferenças de cada amostra. Para mais detalhes em relação aos circuitos responsáveis por estas tarefas, consultar Corrêa (2015).

3. RESULTADOS E DISCUSSÃO

A arquitetura proposta foi sintetizada para um dispositivo FPGA Altera Stratix V 5SGXEA7H3F35C3 com uso da ferramenta Quartus II 14.1. Nestas condições, a arquitetura fez uso de 532 ALUTs e 1247 DLRs e atingiu uma frequência máxima de 459 MHz. Em termos de taxa de processamento, a arquitetura é capaz de processar vídeos de resolução HD 1080p a 94 qps (quadros por segundo) e WQXGA a 33 qps quando operando em modo HEVC; e vídeos 4K UHD a 45 qps quando operando em modo H.264/AVC. O corte nos modos de predição causou um impacto na qualidade de codificação de 46,63% BD-Rate nos experimentos realizados com codificação “all intra”.

O trabalho de Kalali (2012) suporta 16 modos angulares para blocos 4×4 e 33 modos para blocos 8×8, porém não implementa os filtros e nem os modos DC e Planar. Para a tecnologia utilizada FPGA Xilinx Virtex 6, o uso de recursos foi de 2381 LUTs e 849 DFFs, e a frequência máxima foi de 150 MHz. Nestas condições está arquitetura se limita a processar vídeos HD 1080p a 30 qps. O impacto em BD-Rate não foi informado pelos autores.

O trabalho de Liu (2013) suporta todas as combinações de modos de predição e tamanho de bloco. Para tecnologia utilizada TSMC 65 nm, o uso de recursos foi de 77K gates e a frequência máxima foi de 600 MHz. Nestas condições esta arquitetura se limita a processar vídeos HD 1080p a 30 qps.

A Tabela 1 expõe estas informações lado a lado.

Tabela 1 – Resultados de Síntese e Performance

	Este Trabalho	Kalali (2012)	Liu (2013)
Tecnologia	Altera Stratix V	Xilinx Virtex-6	TSMC 65 nm
Área	532 ALUTs, 1247 DLRs	2381 LUTs, 849 DFFs	77000 gates
Frequência	459 MHz	150 MHz	600 MHz
Performance	HD 1080p, 94 qps	HD 1080p, 30 qps	HD 1080p, 30 qps
BD-Rate	46,63%	Não Mencionado	-
Padrão	H.264/AVC, HEVC	HEVC	HEVC

4. CONCLUSÕES

Este trabalho apresentou uma solução de baixo custo para a predição intraquadro multipadrão H.264/AVC e HEVC capaz de processar em tempo real vídeos de alta resolução. Os resultados apresentados são competitivos quando comparados aos já publicados na literatura, além deste trabalho apresentar a primeira arquitetura multipadrão.

Como trabalho futuro, planeja-se o desenvolvimento de arquiteturas para predição intraquadro do padrão com suporte à uma maior quantidade de modos de predição e tamanhos de bloco, mantendo desempenho para vídeos UHD.

5. REFERÊNCIAS BIBLIOGRÁFICAS

CORRÊA, M.; PORTO, M.; ZATT, B.; AGOSTINI, L. A Low-Area and High-Throughput Intra Prediction Architecture for a Multi-Standard HEVC and H.264/AVC Video Encoder. In: 28th Symposium on Integrated Circuits and Systems Design, 2015. Salvador. Proceedings of SBCCI, ACM, 2015.

KALALI, E.; ADIBELLI, Y.; HAMZAOGLU, I. A High Performance and Low Energy Intra Prediction Hardware for High Efficiency Video Coding. In: 22nd International Conference on Field Programmable Logic and Applications, 2012. Oslo. Proceedings of FPL, IEEE, 2012. p.719-722.

LIU, C.; SHEN, W.; MA, T.; FAN, Y. A Highly Pipelined VLSI Architecture for All Modes and Block Size Intra Prediction in HEVC Encoder. In: IEEE 10th International Conference ASIC, 2013. Shenzhen. Proceedings of ASICON, IEEE, 2013. p.1-4.

RICHARDSON, I. G. **H.264 and MPEG-4 Video Compression**, 1.ed. Wiley, 2003. 320p.