

ARQUITETURA DE HARDWARE EFICIENTE PARA A ESTIMAÇÃO DE MOVIMENTO FRACIONÁRIA SIMPLIFICADA DO PADRÃO HEVC

HENRIQUE MAICH¹; LUAN AUDIBERT¹; VLADIMIR AFONSO²; LUCIANO AGOSTINI¹; BRUNO ZATT¹; MARCELO PORTO¹

¹Grupo de Arquiteturas e Circuitos Integrados – Universidade Federal de Pelotas – {hdamaich, lpaudibert, agostini, zatt, porto}@inf.ufpel.edu.br

²IF Sul- Instituto Federal Sul-rio-grandense – vafonso@inf.ufpel.edu.br

1. INTRODUÇÃO

Com a evolução dos dispositivos móveis e a grande exploração comercial por parte da indústria, os fabricantes adicionaram novas funcionalidades para manter a vantagem competitiva. Uma destas funções é a capacidade de gravar e reproduzir vídeos digitais em dispositivos portáteis, como celulares (*smartphones*) e *tablets*. Manter uma boa qualidade nos vídeos, mesmo considerando as restrições dos dispositivos portáteis, está diretamente relacionado com o método de compressão utilizado nestes vídeos. Normalmente, a compressão pode ser realizada de duas formas: a) sem perdas, mantendo a mesma qualidade na imagem; b) com perdas. A codificação com perdas, normalmente, obtém melhores taxas de compressão e é feita baseada em um padrão de codificação de vídeo. O padrão que atualmente obtém as melhores taxas de compressão, dado um determinado valor de qualidade de imagem, é o *High Efficiency Video Coding* (HEVC) (ITU, 2013).

O HEVC utiliza diversas técnicas para buscar e reduzir as redundâncias presentes nos vídeos. Uma dessas técnicas é a Estimação de Movimento (ME – *Motion Estimation*), responsável pela maior parte da compressão obtida com os padrões atuais (PURI, 2004). A ME divide os quadros que compõem o vídeo em blocos menores, chamados de PUs (*Prediction Units*), e tem por objetivo reduzir as redundâncias temporais presentes entre os quadros codificados (SULLIVAN, 2012).

Normalmente a ME pode ser dividida em duas etapas, a IME (*Integer Motion Estimation*) e a FME (*Fractional Motion Estimation*). A IME é realizada buscando o melhor casamento de cada bloco de vídeo a ser codificado com blocos processados anteriormente, pertencentes a um quadro já processado, chamado quadro de referência. A FME, por sua vez, é um refinamento aplicado após a IME. Na FME são gerados novos blocos em posições fracionárias, ao redor do bloco em posições inteiras, e é realizada uma nova busca com o objetivo de obter um melhor resultado em relação a IME.

Como este processo para encontrar as redundâncias presentes nos vídeos é altamente complexo, a busca por simplificações no processo de codificação é altamente desejável. Diversas aplicações precisam de hardware específico para realizar tarefas complexas, pois somente desta forma conseguem obter as taxas de processamento exigidas ao custo de um menor consumo de potência. O foco deste trabalho é desenvolver uma arquitetura de hardware para a FME capaz de atingir taxas de processamento para altas resoluções de vídeo, como UHD 4k (3840x2160 pixels) e 8k (7680x4320 pixels), com baixo consumo de área e potência.

2. METODOLOGIA

O padrão de codificação de vídeo HEVC permite, por definição, que 24 tamanhos de blocos (PUs) sejam testados durante a ME, podendo variar de 8x4 ou 4x8 a 64x64 (SULLIVAN, 2012). Porém, este número de tamanhos diferentes de blocos está diretamente relacionado à complexidade do codificador. Tendo isso em mente, testes foram realizados para verificar o uso de cada um desses tamanhos de bloco, para estipular estratégias visando diminuir a complexidade do codificador. Os experimentos foram realizados utilizando o software de referência do HEVC, o HM

(HEVC *Model*) (FRAUNHOFER, 2014), aplicando as condições comuns de testes (CTC – *Common Test Conditions*) (BOSSSEN, 2012).

Nas CTC estão definidas 24 seqüências de vídeo que devem ser utilizadas em condições específicas para a verificação de qualquer experimento realizado para o HEVC. O primeiro experimento realizado foi uma contagem de quais os tamanhos de PUs eram os mais utilizados no HEVC. Os resultados deste experimento são apresentados na Figura 1. Nos resultados mostrados nesse experimento o percentual de seleção dos blocos foi ponderado pelo tamanho do respectivo bloco, ou seja, um bloco 16x16 tem peso 4 em comparação com o bloco 8x8, por ser mais representativo na imagem.

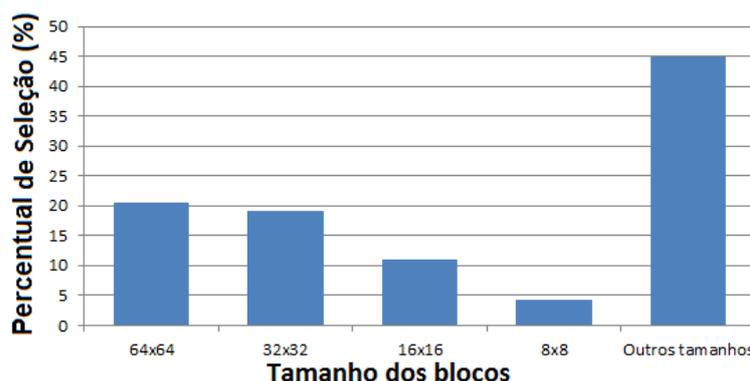


Figura 1 – Percentual de seleção das PUs no HEVC.

É possível verificar que o tamanho mais representativo para a imagem é o 64x64, seguido pelo tamanho 32x32. Baseado nesse resultado, uma boa estratégia é utilizar apenas blocos de tamanho 64x64 na ME. O uso de uma estratégia deste tipo traz uma grande redução no tempo de codificação, uma vez que apenas um tamanho de PU é utilizado na ME ao invés de 24 tamanhos diferentes. Porém, a utilização de apenas um tamanho de bloco na ME implica em perda de qualidade de imagem e, por isso, um segundo experimento foi realizado. Através do segundo experimento foi verificada a degradação na qualidade da imagem considerando alguns cenários em que poucos tamanhos de PUs devem ser testados na ME. Os resultados desse segundo experimento são apresentados na Tabela 1.

Tabela 1 – Resultados de acréscimo no BD-Rate e redução de tempo de codificação.

Tamanho do Bloco	BD-Rate (%)	Redução do Tempo (%)
Somente 64x64	579,69	87,69
Somente 32x32	79	86,79
Somente 16x16	23,51	85,89
Somente 8x8	21,77	82,56
Quadráticos	4,64	58,51

Com os resultados da Tabela 1 foi possível verificar que utilizando apenas os 4 tamanhos quadráticos (8x8, 16x16, 32x32 e 64x64) é possível reduzir o tempo de codificação em 54% na média e, ainda, ter uma baixa perda em compressão para uma mesma qualidade de imagem, aproximadamente 4,64%, na média utilizando a métrica BD-Rate (BJONTEGAARD, 2008). Maiores valores no BD-Rate implicam em perdas na compressão para manter a mesma qualidade. Apesar da redução no tempo de codificação ser grande considerando usar somente um tamanho de bloco, as perdas em termos de compressão e qualidade da imagem são inaceitáveis. Portanto, utilizar apenas blocos quadráticos foi a estratégia utilizada para o desenvolvimento arquitetural da FME.

3. RESULTADOS E DISCUSSÃO

A arquitetura foi descrita em VHDL (*VHSIC Hardware Description Language*) e foi sintetizada para ASIC (*Application Specific Integrated Circuit*) utilizando a ferramenta Cadence *RTL Encounter*. Para diminuir o consumo de hardware, a arquitetura foi projetada para calcular blocos 8x8 e montar os blocos maiores do que o 8x8 a partir dos blocos calculados. A arquitetura completa da FME pode ser visualizada na Figura 2.

Três tipos de Filtros (*Up, Middle, Down*) utilizados para gerar os valores dos blocos fracionários foram implementados, cada um responsável por uma das equações dos filtros definidas pelo HEVC (ITU, 2013). Como o HEVC permite a utilização de bipredição a partir de 2 quadros de referência (SULLIVAN, 2012), foi necessário duplicar os filtros e realizar uma média com as amostras geradas a partir dos dois quadros de referência.

Para a comparação de qual dos blocos gerados trouxe o melhor resultado da FME, foi adotada a métrica SAD (*Sum of Absolute Differences*) (PURI, 2004). Foram utilizadas 12 árvores de SAD, assim obtendo o valor residual gerado por cada bloco. Esses valores de SAD foram acumulados para os 48 blocos, antes de serem comparados, conforme pode ser observado na Figura 2.

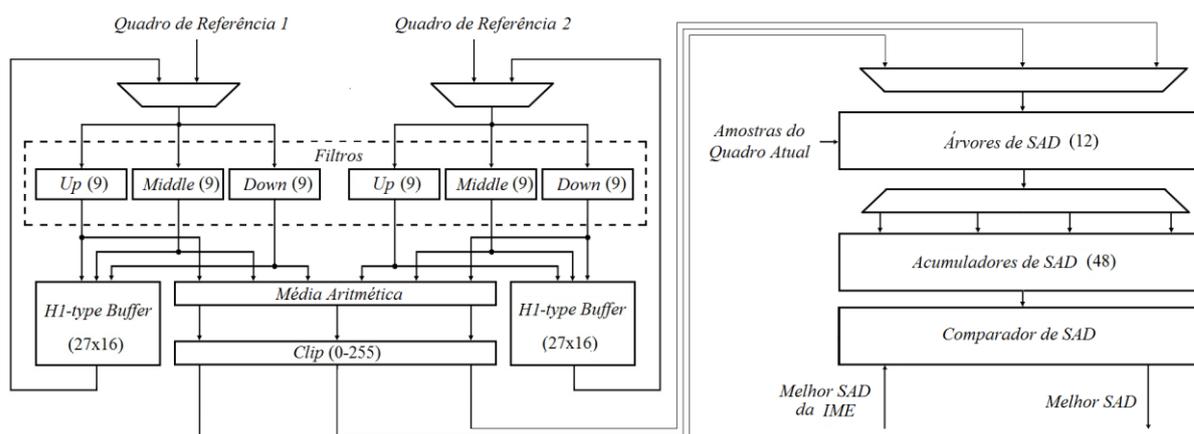


Figura 2 – Modelo da arquitetura desenvolvida.

A arquitetura desenvolvida necessita de 51 ciclos para processar cada bloco 8x8 após a latência, a qual é de 14 ciclos. Desta forma, 1652400 ciclos da arquitetura são necessários para calcular um quadro *Full HD* (1920x1080 pixels), ou seja, 32400 blocos 8x8. Como por definição são necessários ao menos 30 quadros por segundo para ter a sensação de movimento em um vídeo, esse número cresce para 49,58M ciclos, ou uma frequência de 49,58MHz. Para vídeos de resoluções maiores, como 4K ou 8K, esse número é multiplicado por 4 e 16 respectivamente, assim sendo necessárias frequências de 198,3MHz e 793,2MHz.

Tabela 2 – Resultados de síntese e comparação com trabalhos da literatura.

Trabalhos Relacionados	Guo	Huang	He	Arquitetura Desenvolvida
Tecnologia ASIC	90nm	40nm	65nm	45nm
Resultados de Área (gates)	32.496	45.200	1.183k	259.262
Freq. para Full HD@30qps	42.7MHz	50MHz	12MHz	49,58MHz
Freq. para 4K@30qps	171 MHz	200MHz	48MHz	198,3MHz
Freq. para 8K@30qps	não	não	188MHz	793,2MHz

Apenas um trabalho foi encontrado na literatura científica (HE, 2013) que implementa toda a etapa da FME e também é capaz de atingir altas taxas de processamento. Outros trabalhos como (GUO, 2012) e (HUANG, 2013) fazem o processo de interpolação da FME, porém sem realizar a busca entre os blocos calculados. Os dados de síntese para a arquitetura desenvolvida e para os trabalhos relacionados são apresentados na Tabela 2.

É possível verificar que apesar da arquitetura desenvolvida realizar toda a FME, ela necessita de uma baixa frequência de processamento, muito similar a Huang (2013). Porém, a arquitetura desenvolvida também consegue atingir maiores frequências e, assim, processar resolução 8K em tempo real, o que os trabalhos de Guo (2012) e Huang (2013) não conseguem. Quando comparado ao trabalho de He (2013), a arquitetura desenvolvida precisa de maior frequência de processamento para tempo real em uma dada resolução. Contudo, outra diferença importante é que a arquitetura desenvolvida utiliza 78% menos área do que He (2013). Sendo assim, a arquitetura desenvolvida é mais eficiente para ser implementada em dispositivos com limitações em área.

4. CONCLUSÕES

O trabalho propôs uma solução eficaz para a etapa da FME seguindo o padrão de codificação de vídeo HEVC. Para isso, simulações com o HM foram realizadas para verificar quais os tamanhos de blocos são os mais utilizados na ME e, posteriormente, para verificar o impacto em qualidade e tempo de codificação ao realizar simplificações na ME em relação ao que está definido no padrão.

Além disso, baseado nessas análises, uma arquitetura de hardware capaz de processar quadros em altíssima resolução, como 4k e 8k, em tempo real foi desenvolvida. Quando comparado com outros trabalhos da literatura, esta arquitetura obteve altas taxas de processamento com baixo consumo em área, reduzindo em 78% a área consumida pelo melhor trabalho encontrado na literatura.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- BJONTEGAARD, G. Improvements of the BD-PSNR model, **VCEG-AI11**, July 2008.
- BOSSEN, F. **Common Test Conditions and Software Reference Configurations**. Outubro, 2012. Acessado em 1 out. 2013. Online. Disponível em: http://phenix.it-udparis.eu/jct/doc_end_user/current_document.php?id=6469
- FRAUNHOFER. **HEVC Reference Software – HM13.0rc1**. Dec, 2014. Acessado em 1 abril 2014. Online. Disponível em: <http://hevc.hhi.fraunhofer.de>
- GUO, Z.; Zhou, D.; Goto, S. An optimized MC interpolation architecture for HEVC. **IEEE ICASSP**, March 2012.
- HE, G.; et al. A 995Mpixels/s 0.2nJ/pixel fractional motion estimation architecture in HEVC for Ultra-HD, **IEEE A-SSCC**, 2013.
- Huang C.; et al. HEVC Interpolation Filter Architecture for Quad Full HD Decoding, **IEEE VCIP**, 2013.
- ITU - International Telecommunication Union. **Recommendation ITU-T H.265: High efficiency video coding**. Abril, 2013. Acessado em 1 out. 2013. Online. Disponível em: <http://www.itu.int/rec/T-REC-H.265-201304-I>
- PURI, A.; et al. Video Coding Using the H.264/MPEG-4 AVC Compression Standard. **Elsevier Signal Processing: Image Communication**, n. 19, pp. 793–849, 2004.
- SULLIVAN, G. J.; et al. Overview of the High Efficiency Video Coding (HEVC) Standard. **IEEE TCSVT**, Vol. 22, 1649–1668, Dec. 2012.