

## DESENVOLVIMENTO DE ARQUITETURA PARA ESTIMAÇÃO DE MOVIMENTO FRACIONÁRIA DO PADRÃO HEVC

HENRIQUE MAICH<sup>1</sup>; VLADIMIR AFONSO<sup>2</sup>; BRUNO ZATT<sup>1</sup>; LUCIANO AGOSTINI<sup>1</sup>; MARCELO PORTO<sup>1</sup>

<sup>1</sup>Universidade Federal de Pelotas  
{hdamaich, zatt, agostini, porto}@inf.ufpel.edu.br

<sup>2</sup>Instituto Federal Sul-rio-grandense  
vladimir@pelotas.ifsul.edu.br

### 1. INTRODUÇÃO

Com o avanço tecnológico, os dispositivos portáteis, como smartphones e tablets, passaram a permitir a visualização de vídeos digitais em altas resoluções. Contudo, este tipo de vídeo requer o armazenamento e o processamento de uma grande quantidade de dados, o que é extremamente crítico. Essas limitações exigem o emprego de técnicas de codificação de vídeo que permitam a redução do número de dados necessários para representar um vídeo e, conseqüentemente, o espaço necessário para o armazenamento e a largura de banda necessária para a transmissão dos dados.

O estado da arte entre os padrões de codificação de vídeo é chamado HEVC (*High Efficiency Video Coding*) (ITU, 2013) e permite altas taxas de compressão, com elevada qualidade de imagem. Os padrões de codificação de vídeo, como o HEVC, utilizam várias etapas de codificação e uma das etapas mais importantes é a predição interquadros.

A predição interquadros é composta pelos módulos da Estimação de Movimento (ME – *Motion Estimation*) e da Compensação de Movimento (MC – *Motion Compensation*) e explora a redundância que os quadros temporalmente próximos costumam apresentar. Desta forma, é possível diminuir drasticamente os dados para representar os quadros, sendo apenas necessário armazenar a diferença entre o quadro que está sendo codificado e quadros já processados, além de vetores de movimento. Para a aplicação das técnicas de codificação, como a predição interquadros, os quadros são particionados em blocos menores. No padrão HEVC, os blocos utilizados na predição interquadros são chamados de Unidades de Predição (PUs) e podem apresentar tamanhos que variam de 8x4 ou 4x8 até 64x64 amostras (SULLIVAN, 2012). A ME realiza, primeiramente, uma busca apenas em posições inteiras de pixel, a chamada Estimação de Movimento Inteira (IME). Em seguida, uma técnica de refinamento dos resultados chamada de Estimação de Movimento Fracionária (FME – *Fractional Motion Estimation*) pode ser aplicada. A FME é responsável pela geração de amostras de luminância em subpixel, realizada por uma etapa de interpolação, e por uma nova busca utilizando as amostras geradas. O melhor resultado da busca realizada pela FME é comparado ao resultado da IME com o objetivo de definir o vetor de movimento.

A ME é a etapa responsável pela maior parte dos ganhos em termos de compressão (PURI, 2004). Além disso, de acordo com Zhao (2011), a ME é a etapa com maior custo em tempo de processamento, correspondendo de 60 a 90% do tempo total de codificação. Desta forma, o desenvolvimento de arquiteturas para a FME é muito importante, uma vez que soluções em software não são capazes de realizar a codificação de vídeos em altas resoluções em tempo real, principalmente, quando se considera dispositivos alimentados por bateria, os quais apresentam severas restrições em termos de consumo de energia e performance.

Este trabalho apresenta uma arquitetura dedicada para a etapa de interpolação da FME. A arquitetura desenvolvida utiliza PUs de tamanho fixo 16x16 e tem como alvo o processamento em tempo real de vídeos com resolução QFHD (3840x2160) com até 60 quadros por segundo (qps), considerando predição bidirecional com dois quadros de referência.

## 2. METODOLOGIA

No padrão HEVC existem 24 tamanhos diferentes de PU, considerando a predição interquadros. Todos estes tamanhos são avaliados durante a codificação com o objetivo de verificar qual tamanho de PU traz melhores resultados em termos de compressão e qualidade da imagem. Dessa forma, se for utilizado apenas um tamanho de PU na predição interquadros, o esforço computacional para o processamento é drasticamente reduzido ao custo de uma redução na compressão e na qualidade. Para reduzir os efeitos negativos de fixar o tamanho das PUs, algumas avaliações utilizando o software de referência do HEVC (FRAUNHOFER, 2012) foram realizadas. As avaliações foram conduzidas com as condições de testes definidas para o padrão no documento (BOSSSEN, 2012) e tiveram como objetivos verificar quais tamanhos de PUs eram mais selecionados e os efeitos de fixar o tamanho das PUs. Os resultados quantitativos com o percentual de seleções de cada tamanho de PU são apresentados na Figura 1.

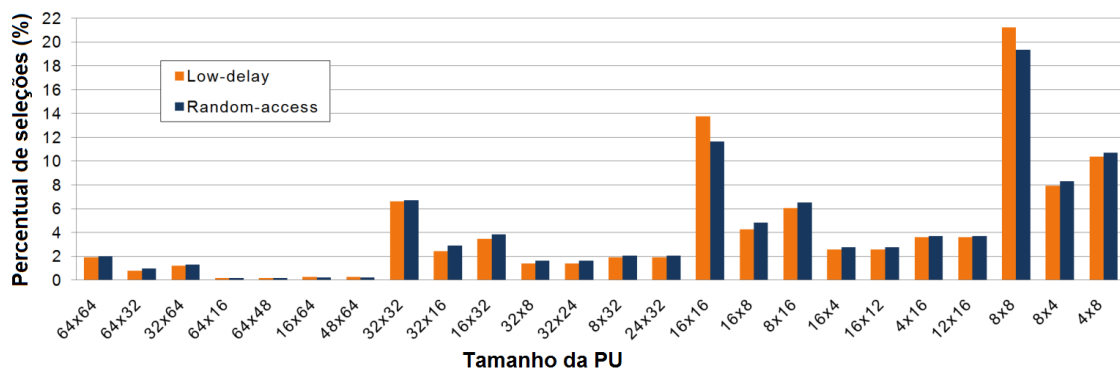


Figura 1 - Percentual de seleções de cada tamanho de PU em média.

Como pode ser observado na Figura 1, os resultados de seleções das PUs são apresentados para cada tamanho de PU, considerando as configurações *Low-Delay* e *Random-Access* (SULLIVAN, 20102). É possível verificar que o tamanho de PU mais selecionado é o 8x8, seguido pelo tamanho 16x16. Estes dois tamanhos foram escolhidos para um novo experimento visando identificar o impacto nas taxas de compressão e qualidade fixando o tamanho da PU. A métrica utilizada para esse novo experimento foi o BD-Rate, uma medida que relaciona a compressão com a qualidade (BJONTEGAARD, 2008). A Tabela 1 mostra o aumento em percentual do Y-BD-Rate fixando o tamanho das PUs em 8x8 ou em 16x16, comparados com uma codificação utilizando todos os tamanhos de PUs. Os valores menores, que representam melhor compressão com mesma qualidade de imagem, são obtidos quando o tamanho da PU é fixado em 16x16. Baseado nos resultados das avaliações, o tamanho 16x16 foi o escolhido para o desenvolvimento da arquitetura.

Tabela 1 - Aumento em percentual do Y-BD-Rate fixando o tamanho das PUs.

Classes de Sequências	Low-delay		Random-access	
	PU 8x8	PU 16x16	PU 8x8	PU 16x16
<b>Classe A – 2560x1600</b>	–	–	38,21	19,06
<b>Classe B – 1920x1080</b>	28,69	21,26	25,42	23,82
<b>Classe C – 832x480</b>	18,93	20,13	15,95	17,16
<b>Classe D – 416x240</b>	14,90	22,28	11,62	17,70
<b>Classe E – 1280x720</b>	33,89	18,05	–	–
<b>Classe F – diversas</b>	21,52	26,85	16,21	19,04
<b>Média</b>	<b>23,59</b>	<b>21,71</b>	<b>21,48</b>	<b>19,36</b>

A arquitetura desenvolvida foi baseada nos algoritmos para interpolação das amostras de luminância definidos pelo padrão HEVC (ITU, 2013). A partir de uma análise dos algoritmos foram adotadas algumas estratégias de otimização visando reduzir o tamanho do hardware desenvolvido e melhorar o seu desempenho em termos de taxa de processamento. As técnicas utilizadas permitiram substituir as multiplicações convencionais, presentes nos algoritmos, por operações de somas e deslocamentos e, também, o compartilhamento de subexpressões, como apresentado no trabalho (AFONSO, 2013).

Para a interpolação das amostras foram desenvolvidos três tipos de filtros, chamados de *Up*, *Middle* e *Down*, sendo que a unidade de processamento foi desenvolvida com três conjuntos destes filtros para manter o paralelismo desejado com a arquitetura. Cada um dos filtros é responsável por gerar uma posição específica a partir das amostras em posições inteiras armazenadas em um *buffer* de entrada. Um multiplexador faz a seleção das amostras na entrada dos filtros. Ainda foram utilizados três *buffers* de saída para o armazenamento das amostras geradas. Estes *buffers* foram chamados de Horizontal, Vertical e Diagonal, e cada um dos *buffers* armazena amostras fracionárias de acordo com as suas posições em relação às amostras inteiras. Cada um dos *buffers* necessita armazenar um número específico de amostras, como pode ser observado na Figura 2.

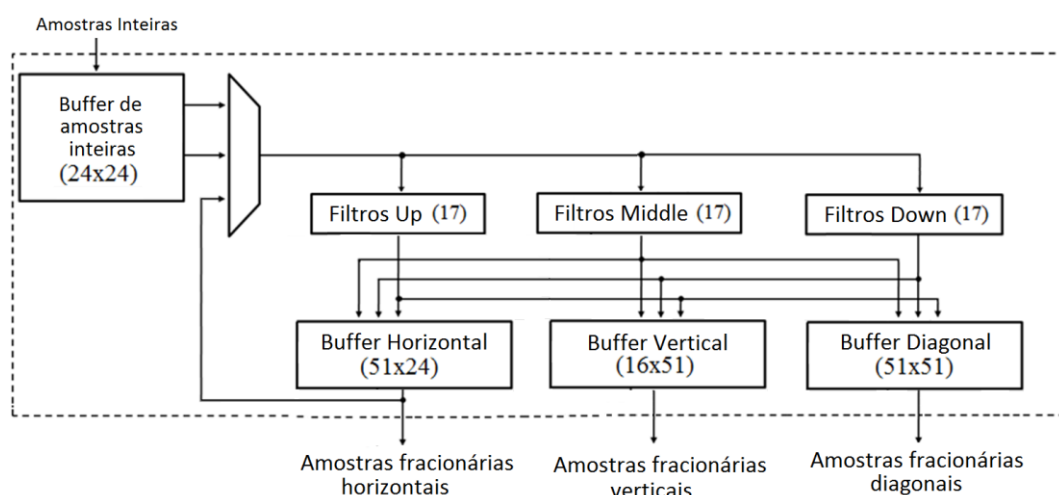


Figura 2 – Visão em alto nível da arquitetura desenvolvida.

### 3. RESULTADOS E DISCUSSÃO

Os resultados de síntese foram obtidos utilizando a ferramenta Quartus II da Altera e consideraram o dispositivo FPGA Altera Stratix III EP3SL200F1517C2. A

Tabela 2 mostra os resultados de hardware para a arquitetura desenvolvida, usando filtros com quatro estágios de *pipeline*.

A arquitetura desenvolvida leva 91 ciclos para interpolar um bloco 16x16 inteiro e, como pode ser observado na Tabela 2, esta arquitetura foi capaz de processar até 60 qps em resolução QFHD (3840x2160) e 240 qps em resolução Full HD (1920x1080), considerando predição bidirecional.

Tabela 2 - Resultados de síntese da arquitetura desenvolvida.

ALUTs Combinacionais	Total de Registradores	Frequência Máxima	QPS máximo para resolução 1080p (1920x1080)	QPS máximo para resolução QFHD (3840x2160)
8.744	57.859	354,9 MHz	240	60

Os resultados da arquitetura desenvolvida foram comparados com os resultados de dois trabalhos relacionados (AFONSO, 2013)(GUO, 2013), que também apresentam arquiteturas para interpolação de amostras de luminância do padrão HEVC. A arquitetura desenvolvida foi a única capaz de atingir 60 qps em resolução QFHD (3840x2160) com predição bidirecional.

#### 4. CONCLUSÕES

Este trabalho apresentou uma solução arquitetural para a etapa de interpolação da FME segundo o padrão HEVC. Diversas avaliações utilizando o software de referência do HEVC foram realizadas, permitindo que fossem conhecidos os tamanhos de PUs mais selecionados e o impacto na codificação de vídeo quando apenas um tamanho de PU é utilizado na predição interquadros. Como a redução do esforço computacional é expressiva quando apenas um tamanho de PU é utilizado, as perdas em termos de compressão e qualidade são aceitáveis, principalmente quando o foco está em dispositivos alimentados por bateria. Baseado nos resultados das avaliações foi desenvolvida uma arquitetura para tamanho de bloco fixo 16x16. Esta arquitetura foi capaz de interpolar 60 qps em vídeos com resolução QFHD (3840x2160 pixels) considerando predição bidirecional, obtendo melhor desempenho que todos os trabalhos encontrados na literatura.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

- AFONSO, V.; Maich, H; Agostini, L.; Franco, D. Low cost and high throughput FME interpolation for the HEVC emerging video coding standard. **4th LASCAS**, 2013.
- BJONTEGAARD, G. Improvements of the BD-PSNR model, **VCEG-A11**, July 2008.
- BOSSSEN, 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](http://phenix.it-udparis.eu/jct/doc_end_user/current_document.php?id=6469)
- FRAUNHOFER. **HEVC Reference Software – HM9.2rc1**. Dec, 2012. Acessado em 1 out. 2013. 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.
- 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.
- ZHAO, Z.; Liang, P. A Statistical Analysis of H.264/AVC FME Mode Reduction. **IEEE TCSVT**, Vol. 21, No. 1, 53-61, Jan. 2011.