

## DESENVOLVIMENTO DE UMA ARQUITETURA MULTITAMANHO DAS IDCTS PARA O DECODIFICADOR DE VÍDEOS HEVC

Ruhan Conceição<sup>1</sup>; José Cláudio de Souza Jr.<sup>1</sup>; Bruno Zatt<sup>1</sup>, Marcelo Porto<sup>1</sup>,  
 Luciano Volcan Agostini<sup>1</sup>

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

### 1. INTRODUÇÃO

Atualmente, há uma crescente popularização de dispositivos capazes de manipular vídeos digitais, como por exemplo câmeras digitais, smartphones, tablets, etc., popularizando o uso deste tipo de vídeo. Em contrapartida, há alguns desafios associados à manipulação destes vídeos, entre estes podemos destacar a quantidade de informação necessária para representa-los. Por exemplo, um vídeo em FullHD (1920 x 1080 pixels), com duração de duas horas, com trinta quadros por segundo; necessitaria de aproximadamente 1,2 terabytes em sua representação. Por outro lado, um vídeo com as mesmas características, codificado com o padrão HEVC – *High Efficiency Video Coding* (ITU-T, 2013), necessitaria apenas de cerca de 4 gigabytes em sua representação. Assim, podemos ver a relevância do processo de codificação e compressão do vídeo.

Um codificador de vídeos é composto por uma sequência de passos, onde cada um dos mesmos é responsável por cada etapa do processo de codificação. Dentre estes, as transformadas ocupam uma posição proeminente. Basicamente, neste módulo os coeficientes são transformados do domínio do espaço para o domínio das frequências. Desta forma a energia do bloco estará concentrada em poucos coeficientes e sendo assim, é aplicada a quantização (gerando matrizes esparsas) e, finalmente, a codificação de entropia.

Dentre os algoritmos de transformação, o HEVC estipula a Transformada Discreta dos Cossenos (*Discrete Cosine Transform* – DCT). Uma das inovações propostas pelo HEVC destaca-se o uso de tamanhos maiores de transformadas, como 32x32 e 16x16, no processo de codificação. O uso destes tamanhos aumenta as taxas de compressão obtidas pelo codificador, entretanto a complexidade computacional do mesmo também é acrescida.

Análogo ao processo de codificação há o processo de decodificação, onde as informações comprimidas são convertidas de uma forma binária para uma forma audiovisual. Dentre os módulos de decodificação está presente a DCT inversa (*Inverse Discrete Cosine Transform* – IDCT), onde é efetuado o cálculo inverso da DCT, transformando os coeficientes transformados do domínio das frequências para o domínio espacial.

Tendo em vista a alta complexidade computacional empregada no processo de codificação – e decodificação – de vídeos, faz-se necessário o desenvolvimento de arquiteturas dedicadas para executar o processo de codificação. Isto deve-se ao fato de que microcomputadores pessoais utilizam abordagens de propósito geral e que geralmente não estão aptos a de/codificar um vídeo em tempo real; ou seja, processar trinta quatros por segundo.

Este trabalho tem por objetivo apresentar o desenvolvimento de uma arquitetura capaz de processar todos os tamanhos de IDCT em uma única estrutura de hardware, utilizando a arquitetura de maior tamanho (32x32) como arquitetura base.

## 2. METODOLOGIA

A fim de desenvolver a arquitetura proposta, primeiramente avaliou-se o software de referência, o HEVC Model (HM) (JCT-VC, 2012). Desta forma, obteve-se os algoritmos utilizados para o desenvolvimento do trabalho. A partir das avaliações realizadas, foram desenvolvidos softwares específicos para cada tamanho de IDCT estipulada pelo padrão. Desta forma, pode-se avaliar a melhor forma de integração entre os diferentes tamanhos das IDCT's. Após, foi gerado um código capaz de processar todos os tamanhos de IDCT's, descrito de forma a tornar fácil uma tradução para linguagem de descrição de hardware.

Tendo em vista o grande número de multiplicadores com constantes fixas para todo o processo, aplicou-se a técnica de substituição de multiplicadores por somas e deslocamento. Multiplicadores completos (sem nenhuma forma de otimização), são extremamente custosos em hardware devido ao elevado número de portas lógicas necessárias para implementá-los (CARRO, 2001). Através desta substituição, observou-se uma quantidade relevante de operações compartilhadas entre as equações de todas as IDCT's, assim a técnica de compartilhamento de sub-expressões foi efetuada a fim de economizar recursos de hardware (CONCEIÇÃO, 2012).

A partir do software otimizado, foi desenvolvida uma arquitetura descrita em VHDL e sintetizada em um dispositivo FPGA Altera Stratix V 5SGXMABN3F45I4 utilizando o software Altera Quartus II 12.1. Além disso, as arquiteturas de tamanho único foram sintetizadas a fim de mostrar os ganhos em termos de consumo de hardware obtido com a solução configurável.

Os softwares gerados a partir do código de referência e as arquiteturas foram validadas comparando os resultados gerados pelos mesmos a partir de amostras extraídas de um processo real de codificação de vídeo utilizando o HEVC Model, para isto o software de referência foi modificado a fim de gravar em arquivos de texto as amostras de entrada e saída do processo das IDCT's.

## 3. RESULTADOS E DISCUSSÃO

Alguns trabalhos relacionados foram encontrados na literatura. Entretanto, devido ao fato das implementações terem sido feitas em tecnologias diferentes, uma comparação justa com os demais trabalhos acaba se tornando uma tarefa difícil. Dentre os trabalhos relacionados, podemos destacar (SHEN, 2011) e (ZHU, 2012).

Assim como mencionado anteriormente, neste trabalho é apresentado uma arquitetura para o cálculo de todas as IDCT 2-D estipuladas pelo HEVC sem o uso de multiplicadores e com compartilhamento de sub-expressões em uma única estrutura. A Tabela 1 apresenta resultados de síntese das arquiteturas desenvolvidas (tamanho único e a multitamanho), como o número de Módulos de Lógica Adaptativa (*Adaptative Logic Module - ALM*), número de registradores e frequência atingida. Além disso, é apresentado a soma dos recursos de hardware utilizados pelas arquiteturas sintetizadas de forma separa a fim de comparar os ganhos da arquitetura proposta.

Tendo em vista as frequências obtidas pelas arquiteturas implementadas, a Tabela 1 também apresenta resultados de desempenho atingidos pelo trabalho para a resolução Full HD (1920x1080) em *qps* (quadros por segundo). Afim de obter a quantidade de amostras necessárias para processar vídeos em tempo real consideramos o espaço de cores YCbCr utilizando subamostragem 4:2:0 (RICHARDSON, 2003), onde para cada 4 amostras de luminância (Y), há uma de

crominância azul (Cb) e uma de crominância vermelha (Cr). Assim, para 1920 x 1080 pixels, é necessário que a arquitetura processe 2073600 (1920 x 1080) amostras de luminância mais 1036800 (1920x1080/4 + 1920x1080/4) amostras de crominância, totalizando 3110400 amostras por quadro. A partir destes cálculos, sabendo o número de amostras que a arquitetura processa por segundo, pode-se obter o número de quadros processados por segundo.

Os resultados de desempenho apresentados na Tabela 1 para a arquitetura multitamanho definem o desempenho máximo obtido, quando a arquitetura processa blocos 32x32, e o resultado mínimo, quando a arquitetura processa blocos 4x4. Isto deve-se ao fato da abordagem seguida no projeto, onde a taxa de amostragem (vazão de dados) da configuração é proporcional ao tamanho de bloco a ser processado. Assim, os maiores tamanhos apresentam melhores taxas de processamento.

Tamanho	ALMs	Registradores	Frequência	Full HD qps
4x4	425	391	116,43 MHz	148
8x8	1.478	1.373	94,49 MHz	240
16x16	4.755	4.793	76,76 MHz	388
32x32	16.594	18.234	68,12 MHz	696
Total	23.252	24.791	-	
Multitamanho	17.340	18.409	63,25 MHz	60 – 648

Tabela 1. Resultados de síntese obtidos

Analisando a tabela 1, podemos perceber que a arquitetura proposta consome 74,57% ALMs comparado com as arquiteturas implementadas de forma separada. Considerando o número registradores, a arquitetura proposta consome apenas 74,25%. Ou seja, numa visão geral vemos que a solução proposta implica em um consumo de 25% dos recursos de hardware fornecidos pelo dispositivo FPGA. Além disso, pode-se verificar que a arquitetura é capaz de processar vídeos em resolução Full HD em qualquer uma das configurações (4x4, 8x8, 16x16 e 32x32).

#### 4. CONCLUSÕES

Este trabalho apresentou o desenvolvimento de uma arquitetura otimizada dedicada para o processamento de todos os tamanhos das transformadas discretas dos cossenos inversas especificadas pelo padrão de codificação de vídeo HEVC.

Resultados mostraram que a arquitetura conseguiu atingir os objetivos propostos, atingindo um desempenho suficiente para codificar vídeos Full HD em tempo real. Além disso, através da síntese das arquiteturas de único tamanho, concluiu-se que a solução multitamanho apresentou cerca de 25% de ganho em termos de consumo de hardware.

Desta forma a arquitetura aqui apresentada pode ser integrada em um decodificador de vídeo que segue as especificações do padrão de codificação de vídeo HEVC.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

CONCEIÇÃO, R. Á; JESKE, R; MATTOS, J; AGOSTINI, L. Software para Otimização do hardware da Transformada Discreta do Cosseno 1-D de tamanho 32 do padrão emergente em codificação de vídeos HEVC. **Congresso de Iniciação Científica (CIC)**, Pelotas, RS, 2012.

AGOSTINI, Luciano. **Desenvolvimento de Arquiteturas de Alto Desempenho Dedicadas a Compressão de Vídeo Segundo o Padrão H.264/AVC. 2007**. Data de publicação. Tese (Doutorado em Ciência da Computação) – Instituto de Informática, UFRGS, Porto Alegre, RS, 2007.

CARRO, L. **Projeto e Prototipação de Sistemas Digitais**. Porto Alegre: UFRGS, 2001.

RICHARDSON, I. **H.264 and MPEG-4 Video Compression: Video Coding for Next-Generation Multimedia**. Chichester: John Wiley and Sons, 2003.

SHEN, S.; Low cost and high throughput multiplierless design of a 16 point 1-D DCT of the new HEVC video coding standard. **ICME 2012**, Melbourne, 2012.

ZHU, J.; Fully pipelined DCT/IDCT/Hadamard unified transform architecture for HEVC Codec. **ISCAS 2012**, China, 2012.

International Telecommunication Union (ITU). **ITU-T Recommendation H.265: High Efficiency Video Coding, Audiovisual and Multimedia Systems**, April 2013. Online. Disponível em: <http://www.itu.int/rec/T-REC-H.265-201304-I>.

JOINT COLLABORATIVE TEAM ON VIDEO CODING (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 – HM9: High Efficiency Video Coding (HEVC) Test Model 9 Encoder Description. 11th Meeting: Shanghai, CN, 10–19 Outubro 2012.

JOINT COLLABORATIVE TEAM ON VIDEO CODING (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 Proposed editorial improvements for High Efficiency Video Coding (HEVC) text specification draft 9 (SoDIS). 12th Meeting: Geneva, CN, 14–23 Janeiro 2013.