

PROJETO EM HARDWARE COM BAIXO CUSTO E ELEVADA TAXA DE PROCESSAMENTO PARA A TRANSFORMADA DCT 16X16 DO PADRÃO EMERGENTE DE CODIFICAÇÃO DE VÍDEOS HEVC

RICARDO GARCIA JESKE; JOSÉ CLÁUDIO DE SOUZA JR.; RUHAN CONCEIÇÃO; JÚLIO CARLOS BALZANO DE MATTOS; LUCIANO VOLCAN AGOSTINI

UFPel – Universidade Federal de Pelotas
{rgjeske, jcdsouza, radconceicao, julius, agostini}@inf.ufpel.edu.br

1. INTRODUÇÃO

Os vídeos digitais vêm sendo utilizados em um número cada vez mais elevado de dispositivos e utilizando resoluções cada vez mais elevadas. Por isso, o estudo e o aperfeiçoamento de codificadores/decodificadores de vídeos é uma atividade de extrema relevância no cenário atual. Assim, questões como taxa de compressão, qualidade do vídeo, complexidade computacional, tamanho do hardware utilizado e consumo de energia devem ser investigados.

O H.264/AVC (ITU-T, 2005) é o padrão mais atual, em uso, na área de codificação de vídeos (AGOSTINI, 2007). Em janeiro de 2010 foi criado o JCT-VC (*Joint Collaborative Team – Video Coding*), formado por especialistas da ITU-T e do ISO/IEC, para começar o desenvolvimento do novo padrão de codificação, chamado HEVC – *High Efficiency Video Coding* (JCT-CV, 2012). O objetivo do JCT-VC é aumentar em 50% a compressão dos vídeos em relação ao padrão H.264/AVC, sem que haja perdas significativas na qualidade final da imagem. A versão final do HEVC deve ser lançada no início de 2013.

Este artigo apresenta o desenvolvimento em hardware da DCT 2-D 16x16, usada no HEVC. A DCT 2-D do HEVC é composta por duas instâncias da DCT 1-D de 16 pontos. A etapa de transformadas é uma das inovações que estão sendo propostas pelo HEVC, tanto pelo seu tamanho variável (de 4x4 a 32x32) quanto pelo uso de transformadas com maiores dimensões do que as tradicionais 4x4 e 8x8, além de também utilizar transformadas de tamanhos retangulares. Como o HEVC ainda está em fase de desenvolvimento, existem poucas publicações na literatura relacionadas a implementações em hardware para as ferramentas de codificação previstas para este emergente padrão de codificação de vídeo.

2. MATERIAIS E MÉTODOS

Uma transformada genérica 2-D, pode ser definida como apresentado em (1) e (2), onde N é o número de pontos de DCT, $x(n)$ é a entrada na posição (n) da matriz de entrada, e X_k é o coeficiente de saída (AHMED, 2012). Uma vez que este trabalho está focado em uma DCT de 16x16, então $N = 256$.

$$X_k = \sqrt{\frac{2}{N}} \epsilon_k \sum_{n=0}^{N-1} x(n) \cdot \cos\left[\frac{\pi(2n+1)k}{2N}\right] \quad (1)$$

$$\epsilon_k = \begin{cases} 1 & \text{se } k = 0 \\ \sqrt{2} & \\ 1 & \text{caso contrário} \end{cases} \quad (2)$$

A aplicação direta de equações (1) e (2), sem qualquer simplificação, usará 65.536 multiplicações e 65.280 adições para gerar todos os 256 coeficientes transformados. Uma abordagem tradicional para reduzir a complexidade da DCT 2-D é o uso da propriedade da separabilidade. A separabilidade considera que duas DCT 1-D podem ser aplicadas para gerar a DCT 2-D. Portanto, a DCT 2-D é composta de duas etapas subsequentes de transformações DCT 1-D, ligados por um passo de transposição (AHMED, 2012). Este processo é ilustrado na Figura 1. Uma vez que este trabalho está focado na DCT 16x16 do HEVC, duas DCT 1-D com 16 pontos são necessárias.

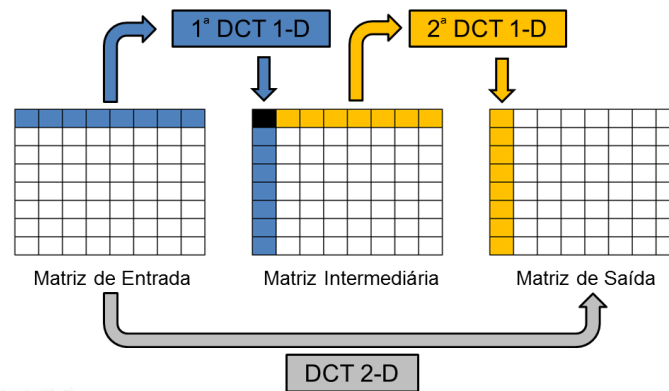


Figura 1. DCT 2-D implementada com separabilidade

O uso da propriedade da separabilidade, reduz o número de operações para 8.192 multiplicações e 7.680 adições. Mas outras simplificações foram propostas neste artigo para reduzir ainda mais o custo do hardware.

Para as otimizações adicionais, todas as equações foram analisadas individualmente e em conjunto, de forma a se obter as melhores sequências de pares de somas com o maior número de reaproveitamentos de operações. Todas as multiplicações foram substituídas por somas e deslocamentos (PARHAMI, 1999) conforme o exemplo a seguir, onde a operação em (1) foi substituída pela operação em (2).

$$X_1 = 72 * a_n \quad (1)$$

$$X_1 = (a_n \lll 6) + (a_n \lll 3) \quad (2)$$

Outra simplificação realizada foi a fatoração e o compartilhamento de subexpressões entre as equações. Cada soma gerada para uma determinada equação é compartilhada entre as demais equações da DCT 1-D.

Considerando todas as otimizações realizadas no algoritmo da DCT 1-D, são necessárias 223 operações de somas ou subtrações e nenhuma multiplicação para gerar o resultado final da DCT 1-D. Como a transposição não utiliza nenhuma operação aritmética, o número total de operações para gerar a DCT 2-D 16x16 inclui 16 cálculos da primeira DCT 1-D (uma para cada linha da matriz de entrada) e 16 cálculos da segunda DCT 1-D (uma para cada coluna da matriz de resultados da primeira DCT 1-D). Assim, o custo total da DCT 2-D é de 7.136 operações de somas ou subtrações, contra as 65.536 multiplicações e 65.280 adições da versão sem separabilidade ou 8.192 multiplicações e 7.680 adições da versão com separabilidade. Considerando o custo elevado das operações dos multiplicadores (no mínimo oito vezes mais custosos que um somador), fica claro que são muito expressivos os ganhos obtidos pelas simplificações realizadas.

Todas as otimizações foram validadas através de avaliações comparativas com o software de referência do padrão HEVC, chamado de HM (HEVC Model). A versão 5.0 do HM foi usada neste trabalho (JCT-VC, 2011). A validação demonstrou que as equações simplificadas eram equivalentes às equações originais. Assim, foi possível iniciar o projeto em hardware para estas equações.

O projeto arquitetural da DCT 1-D de 16 pontos usou como base o algoritmo otimizado e a arquitetura foi descrita em VHDL de forma puramente combinacional e de forma completamente estrutural. As entradas da arquitetura foram definidas em nove bits e as saídas foram definidas com dezoito bits, para manter a precisão do resultado e para garantir que nenhum *overflow* aconteça. Os somadores foram implementados como *ripple carry* de forma estrutural. Os subtratores foram descritos a partir dos somadores, considerando uma das entradas invertida e o *carry-in* com o valor 1.

A arquitetura desenvolvida utilizou 223 operadores de soma ou subtração, com variações de nove a dezenove bits. Nos piores casos existem oito somadores ou subtratores em série.

A otimização do algoritmo foi efetuada com foco na implementação em hardware e, por isso, as multiplicações foram organizadas de tal modo que pudessem ser implementadas através de simples concatenações de zeros aos operandos.

A arquitetura da DCT 2-D 16x16 utiliza duas DCT 1-D de 16 pontos e duas matrizes de registradores para realizar a transposição. O processo de leitura e escrita se dá de forma alternada, enquanto a primeira DCT está escrevendo em uma das matrizes de registradores, a segunda DCT está lendo os dados da outra matriz.

As arquiteturas foram validadas através de comparações entre os resultados gerados pelo software de referência do padrão HEVC (HM 5.0) e os resultados gerados pela arquitetura. Neste caso, o HM foi executado com uma sequência de vídeo e os dados de entrada e de saída do módulo da DCT 2-D 16x16 foram capturados e gravados em arquivos. Assim foi possível estimular a arquitetura com o mesmo conjunto de dados processados pela DCT 2-D 16x16 do HM e foi possível comparar os resultados de saída da arquitetura com os resultados gerado pelo HM.

3. RESULTADOS E DISCUSSÕES

A arquitetura da DCT 2-D 16x16 foi sintetizada para um dispositivo FPGA EP3SL50F780C2 da família Stratix III da Altera. Os resultados de síntese para as arquiteturas da DCT 1-D e da DCT 2-D são apresentados na Tabela 1.

Tabela 1. Resultados de Síntese

| Arquitetura | ALUTs | Frequência (MHz) |
|-------------|--------|------------------|
| 1-D DCT | 5.168 | 87,60 |
| 2-D DCT | 16.002 | 27,05 |

Stratix III EP3SL50F780C2

Os resultados da Tabela 1 mostram que a DCT 2-D apresentou uma redução na sua frequência de operação máxima quando comparada com a arquitetura da DCT 1-D. Isto já era esperado, uma vez que foram realizados um maior número de operações e com somadores utilizando um número maior de

bits. Mas como a arquitetura é capaz de processar 16 amostras a cada ciclo do *clock*, esta frequência permite taxas de processamento muito elevadas. A arquitetura é capaz de processar 432,79 milhões de amostras do vídeo por segundo. Isso permite o processamento de 139 quadros Full HD (1920x1080 pixels) por segundo ou 34,79 quadros Quad Full HD (3840x2160 pixels) por segundo. Em ambos os casos, a arquitetura é capaz de processar vídeos de elevada resolução em tempo real (30 quadros por segundo).

4. CONCLUSÕES

O objetivo deste trabalho foi gerar uma arquitetura de baixo custo e alto desempenho para a DCT 2-D 16x16 do padrão emergente HEVC. A arquitetura da DCT 2-D foi desenvolvida explorando a propriedade da separabilidade e simplificações foram realizadas para reduzir ainda mais o custo do hardware: (a) multiplicações substituídas por somas e deslocamentos; (b) uso de fatoração para reduzir a largura de bits dos operadores e (c) compartilhamento de subexpressões. O custo total da DCT 2-D com separabilidade e sem as otimizações propostas neste trabalho foi de 8.192 multiplicações e 7.680 adições. Com as otimizações propostas, é possível realizar a DCT 2-D com apenas 7.136 operações de somas ou subtrações. Esse resultado é especialmente relevante se for considerado que o custo de um multiplicador é muito mais elevado do que o custo de um somador ou subtrator.

A arquitetura desenvolvida foi projetada de forma puramente combinacional e é capaz de processar 16 amostras por ciclo de *clock*.

Os resultados de síntese mostram que a arquitetura da DCT 2-D é capaz de atingir uma taxa de processamento superior a 400 milhões de amostras por segundo, sendo capaz de processar vídeos em resoluções muito elevadas, como QFHD (3840x2160 pixels), em tempo real (30 quadros por segundo).

5. REFERÊNCIAS BIBLIOGRÁFICAS

International Telecommunication Union. "ITU-T Recommendation H.264/AVC (03/05): advanced video coding for generic audiovisual services". 2005.

AGOSTINI, L. V.; et al. Design and FPGA Prototyping of a H.264/AVC Main Profile Decoder for HDTV. **Journal of the Brazilian Computer Society**, v. 13, p. 25-36, 2007.

Joint Collaborative Team on Video Coding (JCT-VC). Acessado em 05 de março de 2012. Disponível em: <http://www.itu.int/en/ITU-T/studygroups/com16/video/Pages/jctvc.aspx>

AHMED, A.; SHAHID, M. U.; REHMAN, A. N-Point DCT VLSI Architecture for Emerging HEVC Standard. **Hindawi**. 2012.

PARHAMI, B. **Computer arithmetic: algorithms and hardware designs**. New York, 1999. 1v

Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 – "HM5: High Efficiency Video Coding (HEVC) Test Model 5 Encoder Description", 7th Meeting: Geneva, CH, 21-30 November, 2011.