

PROJETO DE HARDWARE PARA A TRANSFORMADA ROTACIONAL DO PADRÃO EMERGENTE DE CODIFICAÇÃO DE VÍDEO HEVC

HENRIQUE AVILA VIANNA; LUCIANO VOLCAN AGOSTINI

*Universidade Federal de Pelotas – PPGC/Mestrado em Ciência da Computação
Grupo de Arquiteturas e Circuitos Integrados (GACI)
{havianna,agostini}@inf.ufpel.edu.br*

1. INTRODUÇÃO

Os vídeos digitais tornaram-se um importante recurso de entretenimento, comunicação e educação, com presença crescente em nosso cotidiano através da televisão digital, tocadores de DVD e Blu-ray, computadores pessoais, celulares e diversos outros dispositivos. A popularização dos *displays* de alta-definição e de equipamentos domésticos capazes de produzir vídeos em HD tem feito crescer a demanda por resoluções mais elevadas e maior fidelidade de imagem. Entretanto, as redes de comunicação de dados, especialmente as tecnologias sem-fio, continuam sendo um fator limitante para a transmissão de vídeos de alta-definição com alta qualidade (ISO/IEC, 2011), devido ao elevado volume de dados necessário para representar um vídeo em forma digital.

A codificação de vídeo explora o alto grau de similaridade entre as imagens que compõem um vídeo, codificando essas informações de forma a evitar a transmissão e o armazenamento de dados redundantes (AGOSTINI, 2007). Uma codificação mais eficiente permite transmitir ou armazenar mais conteúdo com a mesma qualidade, ou aumentar a resolução ou fidelidade do vídeo utilizando a mesma quantidade de bits. Assim, dada a importância dos vídeos digitais na atualidade, existe um esforço constante da academia e da indústria em desenvolver novos e mais eficientes padrões de codificação.

Em janeiro de 2010 foram iniciados os trabalhos para elaborar um padrão para a próxima geração de codificadores de vídeo. O projeto foi batizado de HEVC (*High Efficiency Video Coding*) e os trabalhos estão sendo coordenados pelo JCT-VC (ITU, 2012). O objetivo do HEVC é obter um ganho significativo de compressão em relação ao atual estado-da-arte, mantendo a mesma qualidade de imagem. Para atingir esse objetivo, estão sendo propostos elementos estruturais mais avançados e novas ferramentas algorítmicas.

A Transformada Rotacional (ROT) é uma das novas ferramentas de codificação propostas para o futuro padrão HEVC. O objetivo desta ferramenta é concentrar melhor as informações nas matrizes de coeficientes resultantes da aplicação da transformada principal (DCT) (MCCANN et al., 2010). A concentração de energia em menos coeficientes favorece a codificação de entropia, além de minimizar a distorção na imagem causada pela quantização.

Devido à elevada complexidade computacional dos diversos algoritmos utilizados na codificação de vídeo, arquiteturas de hardware dedicadas são fundamentais para viabilizar o processamento de vídeos de alta-definição em tempo real e com baixo consumo de energia. O objetivo deste trabalho é apresentar uma arquitetura eficiente para o cálculo da ROT, com foco em vídeos de altíssima resolução, que possa ser utilizada em codificadores do futuro padrão HEVC implementados em hardware.

2. MATERIAL E MÉTODOS

Este trabalho iniciou com a investigação da proposta da Transformada Rotacional (ROT), através dos artigos publicados nos encontros do JCT-VC.

A ROT faz parte do trabalho apresentado por MCCANN et al. (2010) em resposta à chamada de propostas do JCT-VC. A ROT é recomendada como uma transformada adicional, a ser aplicada após a DCT para os blocos de resíduos da predição intra-quadro.

A transformada tem um tamanho fixo de 8x8 amostras. Quando aplicada a blocos maiores do que 8x8, apenas a região de 8x8 amostras do canto superior esquerdo do bloco é processada pela ROT, uma vez que os coeficientes mais significativos da DCT encontram-se nessa região de baixas frequências.

A ROT é um produto de três matrizes, conforme definido na equação (1).

$$W = R_h \cdot M \cdot R_v^T \quad (1)$$

Onde M é um bloco 8x8 de coeficientes da DCT, R_h e R_v são as matrizes de rotação horizontal e vertical respectivamente, e W é o bloco de coeficientes transformados. As matrizes R_h e R_v , por sua vez, são produtos de matrizes de rotação de Givens (GIVENS, 1958), que utilizam senos e cossenos de seis diferentes ângulos para promover uma troca parcial de energia entre as linhas e colunas da matriz de entrada.

A proposta da ROT define quatro conjuntos de ângulos de rotação, que permitem melhor adaptação da transformada para blocos de resíduos com diferentes características. Para cada bloco processado, o codificador deverá calcular as quatro opções de transformada e escolher aquela que resultar no melhor custo taxa-distorção após a quantização. Essa opção é codificada no *bitstream* do vídeo, através do sinal $\text{indexROT} = 1, \dots, 4$ que será utilizado pelo decodificador para reconstruir o bloco original de coeficientes, aplicando a transformada inversa correspondente.

No terceiro encontro do JCT-VC, FERNANDES (2010) apresentou uma proposta de redução de complexidade para a ROT, com uma nova fatoração das matrizes de rotação, que permite o cálculo da transformada através de operações mais simples, utilizando constantes com apenas 5 bits de precisão. Esta técnica foi detalhada nos resultados do experimento apresentado por MA et al. (2011a) e aperfeiçoada em MA et al. (2011b).

Com base nos algoritmos apresentados em MA et al. (2011a) e MA et al. (2011b) extraiu-se as equações da ROT horizontal e vertical. Em seguida, as equações foram simplificadas visando a implementação em hardware com alto desempenho. Como todas as multiplicações envolvem uma constante, foi possível substituir essas operações por somas e deslocamentos, o que reduz o custo do hardware e aumenta seu desempenho em relação a um multiplicador genérico.

A Tabela 1 apresenta as equações originais e simplificadas do primeiro estágio da ROT horizontal. Dependendo do valor de indexROT , a equação correspondente é calculada e o resultado segue para o próximo estágio de cálculos. No total, são utilizadas 72 equações na etapa horizontal da ROT e outras 72 na etapa vertical. Todas as equações compartilham o mesmo formato apresentado na Tabela 1, alterando apenas as variáveis de entrada e as constantes utilizadas nas multiplicações, de acordo com o valor de indexROT .

Tabela 1. Equações originais e simplificadas do primeiro estágio da ROT horizontal, para as diferentes opções de indexROT.

indexROT	equações originais	equações simplificadas
1	$a0 = W0 + (-4 * W1 \gg 5)$	$a0 = W0 - (W1 \gg 3) - 1$
2	$a0 = W0 + (-7 * W1 \gg 5)$	$a0 = W0 - ((W1 \ll 3) - W1 \gg 5) - 1$
3	$a0 = W0 + (2 * W1 \gg 5)$	$a0 = W0 + (W1 \gg 4)$
4	$a0 = W0 + (-1 * W1 \gg 5)$	$a0 = W0 - (W1 \gg 5) - 1$

Após o processo de simplificação das equações, desenvolveu-se uma arquitetura combinacional capaz de processar oito amostras de entrada por ciclo de *clock*. Esta arquitetura recebe como entradas o sinal indexROT e oito amostras de 16 bits, que correspondem a uma linha da matriz de entrada. A arquitetura é dividida em duas partes, uma para a ROT horizontal e outra para a ROT vertical. A ROT horizontal processa a matriz de entrada linha a linha, gerando uma matriz intermediária que é, então, processada coluna a coluna pela ROT vertical, para gerar a matriz transformada final. A matriz intermediária é armazenada em um *buffer* de transposição, composto por 64 registradores de 16 bits. O valor de indexROT, informado apenas no início de cada bloco, é armazenado em registradores dedicados para que o valor correto esteja disponível para as etapas horizontal e vertical. Esses registradores comandam os multiplexadores que selecionam as operações corretas a serem calculadas em cada estágio. O conteúdo dos registradores é atualizado através de um sinal emitido pela parte de controle da arquitetura a cada oito ciclos de *clock*, após o processamento completo de um bloco 8x8.

3. RESULTADOS E DISCUSSÕES

A arquitetura desenvolvida foi descrita em VHDL e sintetizada para um FPGA da família Stratix III, utilizando o software Quartus II. Os resultados da síntese mostram que a arquitetura utilizou 14% dos elementos lógicos do FPGA, sendo 3.647 ALUTs e 1.166 registradores. A frequência máxima de operação atingida foi de 33,17 MHz, o que permite que a arquitetura processe 265,36 milhões de amostras por segundo.

A Tabela 2 apresenta a estimativa de quadros por segundo que a arquitetura é capaz de processar, para diferentes resoluções de vídeo. O processamento de vídeos em tempo real requer uma taxa mínima de 30 quadros por segundo. Pode-se observar nos resultados apresentados na Tabela 2, que a arquitetura atinge o desempenho necessário para o processamento em tempo real de vídeos até a resolução QFHD. Uma versão da arquitetura utilizando uma estratégia de *pipeline* já está sendo desenvolvida, visando atingir a taxa de processamento necessária para a codificação em tempo real também na resolução UHDTV.

Tabela 2. Taxas de processamento, em quadros por segundo, obtidas pela arquitetura desenvolvida, para diferentes resoluções de vídeo.

Resolução	Quadros por segundo
Full HD (1920x1080)	127,9
QHD (2560x1440)	71,9
QFHD (3840x2160)	31,9
UHDTV (7680x4320)	7,9

A arquitetura desenvolvida foi validada utilizando o software de simulação ModelSim. Para a validação foram utilizados dados numéricos de 4.000 blocos de vídeo, processados pelo software de referência do HEVC com a ROT habilitada. Os resultados gerados pela arquitetura foram idênticos aos extraídos do software, para todos os blocos de entrada.

Até onde se tem conhecimento, não há nenhum trabalho publicado na literatura sobre a ROT implementada em hardware. Assim, não foi possível comparar a solução apresentada com outros trabalhos.

4. CONCLUSÕES

Este trabalho apresentou uma arquitetura de hardware dedicada ao cálculo da Transformada Rotacional (ROT), uma nova ferramenta de codificação de vídeo proposta para o padrão emergente HEVC. A arquitetura desenvolvida é capaz de processar vídeos até a resolução QFHD em tempo real, podendo ser empregada em futuros codificadores HEVC implementados em hardware.

A maior contribuição deste trabalho é o fato de ser o primeiro trabalho disponível na literatura a apresentar a implementação da ROT em hardware.

5. REFERÊNCIAS BIBLIOGRÁFICAS

AGOSTINI, L.V. **Desenvolvimento de Arquiteturas de Alto Desempenho Dedicadas à Compressão de Vídeo Segundo o Padrão H.264/AVC**. 2007. 172f. Tese (Doutorado em Computação) – Programa de Pós-Graduação em Computação, Universidade Federal do Rio Grande do Sul.

FERNANDES, F. Low Complexity Rotational Transform. **JCT-VC Meeting**, 3. Guangzhou, 2010. JCTVC-C096.

GIVENS, J.W. Computation of plane unitary rotations transforming a general matrix to triangular form. **SIAM Journal of Applied Mathematics**, v.6, n.1, p.26-50, 1958.

ISO/IEC. Vision, Applications and Requirements for High Efficiency Video Coding (HEVC). **ISO/IEC JTC1/SC29/WG11**, document W11872. 2011.

INTERNATIONAL TELECOMMUNICATION UNION. **Joint Collaborative Team on Video Coding (JCT-VC)**. Acessado em agosto 2012. Online. Disponível em <http://www.itu.int/en/ITU-T/studygroups/com16/video/Pages/jctvc.aspx>

MA, Z. et al. CE 7: Experimental Results for the Rotational Transform. **JCT-VC Meeting**, 6. Torino, 2011a. JCTVC-F294.

MA, Z. et al. Non CE 7: Supplementary Results for the Rotational Transform. **JCT-VC Meeting**, 7. Genebra, 2011b. JCTVC-G591.

MCCANN, K. et al. Samsung's Response to the Call for Proposals on Video Compression Technology. **JCT-VC Meeting**, 1. Dresden, 2010. JCTVC-A124.