

ARQUITETURA EM HARDWARE PARA INTERPOLAÇÃO DE QUARTER-PIXELS SEGUNDO O PADRÃO H.264/AVC

SCHOENKNECHT, Mateus; CORRÊA, Marcel

Grupo de Arquiteturas e Circuitos Integrados – GACI

Universidade Federal de Pelotas – UFPel

AGOSTINI, Luciano

Grupo de Arquiteturas e Circuitos Integrados – GACI

Universidade Federal de Pelotas – UFPel

1 INTRODUÇÃO

O padrão de codificação de vídeo H.264/AVC (*Advanced Video Coding*) (JVT, 2003) é o estado da arte em codificação de vídeo. Ele foi desenvolvido por um grupo de especialistas da ITU-T e ISO/IEC com o intuito de alcançar maiores níveis de compressão quando comparado com padrões anteriores como o MPEG-2.

O H.264/AVC tem alta complexidade computacional (RICHARDSON, 2003), o que torna difícil que soluções em software codifiquem vídeos de alta definição em tempo real (1920x1080 pixels por quadro a 30 quadros por segundo). Por esse motivo, arquiteturas em hardware são a melhor forma de codificar e decodificar vídeos de alta definição em tempo real. Essa investigação sobre soluções em hardware para o padrão H.264/AVC está inclusa nos esforços de pesquisa para a implementação do Sistema Brasileiro de Televisão Digital (SBTVD), que utiliza o H.264/AVC como padrão de codificação de vídeo (ISDTV).

Um vídeo sem compressão tem uma grande quantidade de informação redundante que pode ser explorada e reduzida. Existem basicamente três tipos de redundância: (1) a redundância espacial, que é a similaridade em áreas homogêneas de um mesmo quadro; (2) a redundância temporal, que é a similaridade entre quadros sequenciais, e (3) a redundância entrópica, que é a similaridade na representação dos códigos binários (RICHARDSON, 2003).

A Estimção de Movimento (ME) é o modulo que explora e reduz a redundância temporal de um vídeo durante o processo de codificação, e é também o mais complexo módulo do codificador H.264/AVC. Ela funciona dividindo o quadro atual em diversos macroblocos (16x16 pixels) e, então, buscando nos quadros codificados anteriormente (quadros de referência) pelo macrobloco que é mais similar ao corrente. Depois dessa busca é gerado um vetor de movimento inteiro (VMI) para indicar o deslocamento. O objetivo da estimção de movimento é encontrar o melhor VMI enquanto mantém a complexidade computacional dentro de limites aceitáveis.

Uma característica que contribui para as altas taxas de compressão alcançadas pela ME do H.264/AVC é a possibilidade de gerar vetores de movimento fracionários (VMF) apontando para blocos compostos de sub-pixels (RICHARDSON, 2003). Em outras palavras, um movimento que ocorre de um

quadro para outro não é limitado a posições inteiras. Assim, é integrada à ME inteira alguns módulos para dar suporte à ME fracionária.

Este trabalho propõe uma arquitetura de alto desempenho em hardware para realizar a interpolação das amostras de *quarter-pixel* usadas na Estimção Fracionária de Movimento (EFM). Esta arquitetura será integrada com a busca de *quarter-pixel*, o refinamento de *half-pixel* e a busca inteira da ME para gerar uma arquitetura completa de ME com suporte a precisão fracionária de pixels.

Com base em avaliações realizadas em software, se percebeu que a ME inteira poderia ser limitada a processar somente blocos 8x8, ao invés de utilizar os tamanhos variados de bloco que são suportados pelo padrão H.264/AVC, sem custos expressivos em qualidade. Além disso, foi possível perceber que os tamanhos de bloco iguais ou maiores que 8x8 são os mais usados, justificando a escolha pelo tamanho 8x8 para os blocos da ME. Com estas simplificações, a ME do H.264/AVC passa a ter pequenas perdas em taxa de processamento e em qualidade se comparada ao caso ótimo, mas a complexidade do processo é drasticamente diminuída, justificando plenamente o uso das simplificações. A avaliação foi realizada com 5 vídeos QCIF (176x144 pixels), os quais foram codificados utilizando o software de referência do padrão H.264/AVC e os resultados foram avaliados com base no *bit rate* e no PSNR (*Peak Signal-to-Noise Ratio*) (CORRÊA, 2010).

2 METODOLOGIA

A interpolação dos *quarter-pixels* ocorre após a interpolação dos *half-pixels*. Para gerar a área de busca composta por *quarter-pixels*, a interpolação utiliza o bloco escolhido como melhor casamento na etapa de *half-pixel*. Qualquer tipo de *quarter-pixel* pode ser obtido através da fórmula definida em (1).

$$y = (A + B + 1) \gg 1 \quad (1)$$

Existem três tipos diferentes tipos de *quarter-pixel*: (1) tipo H, que é calculado utilizando as duas amostras horizontais mais próximas; (2) tipo V, que é calculado usando as duas amostras verticais mais próximas e (3) tipo D, que é calculado usando as duas amostras diagonais mais próximas em cada direção. A Figura 1a mostra os *quarter-pixels* do tipo H, a Figura 1b mostra os *quarter-pixels* do tipo V e a Figura 1c mostra os *quarter-pixels* do tipo D.

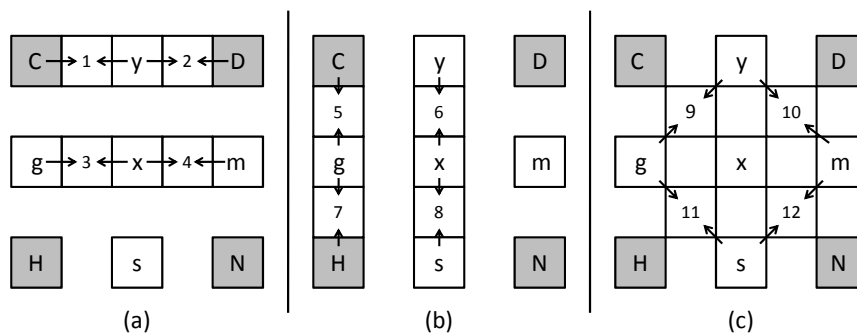


Figura 1 – Tipos de *quarter-pixel*.

A unidade de interpolação de *quarter-pixel* é o módulo responsável pela obtenção das amostras de *quarter-pixel* e pode ser dividido em três partes principais: (1) a linha de filtros, (2) os *buffers* e (3) a unidade de controle.

A linha de filtros é um módulo que interpola uma linha inteira de qualquer tipo de *quarter-pixel* em um único passo. Uma vez que existem 16 amostras de *quarter-pixel* por linha, a linha de filtros deve ter 16 unidades de processamento (UP). A UP é a unidade operativa básica e realiza o filtro bilinear descrito na fórmula (1).

Os *buffers* são necessários para armazenar e deslocar os valores de entrada (uma área de 17x17, composta por um bloco 8x8 de amostras inteiras intercaladas com os *half-pixels* interpolados previamente) e os valores de saída (oito blocos 8x8 compostos de *quarter-pixels*). A Figura 2 mostra a arquitetura da unidade de interpolação cujos sinais de controle foram omitidos para uma melhor visualização.

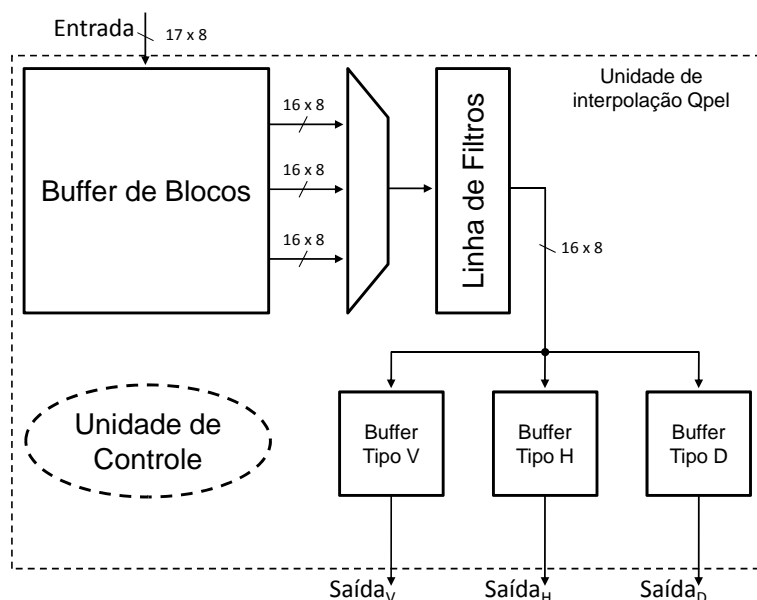


Figura 2 – Arquitetura do Interpolador de *Quarter-Pixel*.

3 RESULTADOS E DISCUSSÕES

A arquitetura proposta foi totalmente descrita em VHDL e sintetizada para um dispositivo FPGA Xilinx Virtex4, utilizando a ferramenta de síntese Xilinx ISE (XILINX, 2010).

Quando sintetizada, a arquitetura atingiu uma frequência máxima de 450 MHz. Nessa frequência, a arquitetura pode processar 53 quadros QHDTV (3840x2048 pixels) por segundo. Contudo, uma vez que a interpolação de *quarter-pixel* é o processo menos complexo de toda a EFM o seu desempenho pode ser limitado por outros processos da EFM.

As frequências mínimas para processar vídeos em tempo real são: 250,67 MHz para QHDTV, 66,1 MHz para HD 1080p (1920x1080), e 29,38 MHz para HD 720p(1280x720).

A parte operativa (linha de filtros) da arquitetura consumiu 176 slices e 324 LUTs de 4 entradas do dispositivo alvo. Os *buffers* foram mapeados como

bancos de registradores e consumiram 801 registradores (8-bits), 289 multiplexadores 4:1 (8-bits) e 512 multiplexadores 2:1 (8-bit) do FPGA alvo.

Essa solução apresenta um bom equilíbrio entre desempenho e custo de hardware, uma vez que é capaz de codificar vídeos de alta definição com uma baixa frequência e consome uma pequena quantidade de recursos em hardware.

4 CONCLUSÕES

Esse trabalho apresentou uma arquitetura de alto desempenho e baixo custo de hardware para o processo de interpolação de *quarter-pixel* do padrão H.264/AVC, tendo como alvo arquiteturas de ME que dão suporte a blocos de tamanho 8x8. Quando sintetizada para o FPGA alvo, a arquitetura alcançou uma elevada taxa de processamento, sendo capaz de processar vídeos de alta resolução (como HD 1080p) em tempo real com uma baixa frequência de operação (66,1 MHz no caso de HD 1080p).

Como trabalhos futuros, planeja-se o desenvolvimento do refinamento de *quarter-pixel* utilizando a arquitetura apresentada neste trabalho, e, então, integrá-la com o refinamento de *half-pixel* e com a ME inteira desenvolvidos em trabalhos paralelos do grupo.

5 REFERÊNCIAS

Brazilian Forum of Digital Television, ISDTV Standard, Draft, Dec, 2006.

Corrêa, M.M.; Schoenknecht, M.T.; Agostini, L.V, "A high performance hardware architecture for the H.264/AVC half-pixel motion estimation refinement," **23rd Symposium on Integrated Circuits and Systems Design (SBCCI)**, 2010 (aceito para publicação).

JVT, JOINT VIDEO TEAM OF ITU-T AND ISO/IEC JTC 1. **Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification** (ITU-T Rec. H.264 or ISO/IEC 14496-10 AVC), 2003.

RICHARDSON, Ian. **H.264 and MPEG-4 Video Compression: Video Coding for the Next-generation Multimedia**. Chichester: Wiley, 2003.

XILINX; "FPGA and CPLD Solutions from Xilinx, Inc." Jul, 2010; <http://www.xilinx.com>.