



ARQUITETURA DE UM MODO DE DECISÃO RÁPIDO BASEADO EM HEURÍSTICA PARA A ESTIMAÇÃO DE MOVIMENTO DO PADRÃO H.264/AVC DE COMPRESSÃO DE VÍDEO

SANCHEZ, Gustavo; DORNELLES, Robson; AGOSTINI, Luciano.

*Grupo de Arquiteturas e Circuitos Integrados – DInfo/IFM/UFPeI
Campus Universitário – Caixa Postal 354 – CEP 96010-900.
{gustavofreitassanchez, dudassdornelles}@gmail.com; agostini@ufpel.edu.br*

1. INTRODUÇÃO

Quanto maior é a qualidade de vídeo necessária em aplicações cotidianas, como a transmissão de televisão digital, maior é a necessidade de comprimir (codificar) o vídeo de maneira mais eficiente, a fim de que o usuário possa receber o vídeo de alta qualidade em tempo real. O mais novo padrão de compressão de vídeo é o H.264/AVC (ITU-T, 2003). O H.264/AVC é capaz de dobrar as taxas de compressão atingidas pelos padrões anteriores (como o MPEG-2) e é utilizado no Sistema Brasileiro de TV Digital (SBTVD).

Comprimir um vídeo significa diminuir a quantidade de bits que é preciso para representá-lo sem perdas significativas em sua qualidade. Para fazer isto, é preciso explorar as redundâncias existentes em um vídeo digital. Existem três tipos de redundância: temporal (semelhanças entre quadros), espacial (semelhanças dentro do próprio quadro) e entrópica (redundância na representação binária dos dados) (RICHARDSON, 2003).

A redundância temporal é explorada, no padrão H.264/AVC, pela Estimação de movimento (ME), um módulo do codificador que compara o bloco que quer codificar com vários blocos de um ou mais quadros de referência (já codificados), escolhendo o bloco com maior semelhança. Esta escolha é realizada utilizando uma métrica de erro, como o SAD (*Soma das Diferenças Absolutas*). Assim, é gerado um vetor que aponta para o bloco escolhido. É trabalho de outro módulo, a Compensação de Movimento, remontar o novo quadro predito, utilizando os vetores gerados pela ME, e calcular o resíduo (diferença entre o quadro predito e o quadro original). Este resíduo é codificado e transmitido, juntamente com os vetores de movimento. Na decodificação, é possível reconstruir o vídeo fazendo novamente a Compensação de Movimento e somando o quadro gerado, ao seu resíduo.

A principal inovação no H.264/AVC, no ponto de vista da estimação de movimento, está na possibilidade de utilização de tamanhos de blocos variáveis para realizar a ME. Ao invés de utilizar um macrobloco (maior tamanho possível de bloco) inteiro na estimação de movimento, o padrão H.264/AVC permite o uso de partições de macroblocos e partições de sub-macroblocos (AGOSTINI, 2005). As partições possuem tamanho 16x16, 8x16, 16x8, 8x8 e as subpartições podem ter tamanho

8x8, 4x8, 8x4 e 4x4. Um dos maiores desafios da ME do padrão H.264/AVC é a tomada de decisão de qual tipo de partição utilizar em cada caso. Este trabalho apresenta uma solução em hardware para esta questão, usando uma heurística para acelerar a tomada de decisão, isto é importante, pois a transmissão de dados deve ser feita o mais rápido possível a fim de que o vídeo possa ser visto em tempo real. Esta solução foi descrita em VHDL e sintetizada para FPGAs, apresentando resultados promissores.

O restante do texto apresenta o trabalho desenvolvido em maiores detalhes. A Seção 2 apresenta os materiais e os métodos utilizados para o desenvolvimento do projeto; a Seção 3 apresenta os resultados da síntese da arquitetura; e finalmente, a Seção 4 conclui este trabalho.

2. MATERIAL E MÉTODOS

Na Estimação de Movimento, o algoritmo de busca com maior complexidade computacional é o *FullSearch* (FS) (RICHARDSON, 2003). O FS busca o melhor casamento para cada bloco 4x4 de um macrobloco 16x16 (que possui 16 blocos 4x4), encontrando sempre o resultado ótimo, já que busca por todas as posições da área de pesquisa. Como um macrobloco com 16x16 amostras do vídeo é formado por 16 blocos de 4x4 amostras, para cada macrobloco serão gerados 16 vetores de blocos 4x4. Utilizando os resultados intermediários, é possível encontrar o melhor casamento também para as partições 8x4, 4x8, 8x8, 16x8, 8x16 e 16x16. Este trabalho propõem a implementação arquitetural de um algoritmo baseado em heurística para a decisão de quais vetores serão utilizados para um macrobloco 16x16. A heurística proposta é baseada na constatação de que existe uma elevada correlação entre vetores de movimento de blocos vizinhos. Então, caso os vetores de blocos 4x4 vizinhos sejam idênticos, os dois blocos podem ser agrupados, gerando um bloco maior. O processo é repetido até que macroblocos 16x16 sejam formados. O uso de blocos maiores, nesse caso, aumenta a taxa de compressão sem reduzir a qualidade do vídeo. A heurística ainda prevê o agrupamento de vetores similares, com pouca diferença. Nesse caso, pode haver perda de qualidade de imagem em favor de uma maior taxa de compressão.

Com a heurística definida, foi elaborado um algoritmo para esta solução e, posteriormente, foi projetada uma arquitetura que recebe os 16 vetores de movimento de 16 blocos 4x4 e decide como ficará a partição do macrobloco. No padrão H,264/AVC, a ME de um macrobloco 16x16 é organizada em um formato de duplo Z (AGOSTINI, 2005), ou seja, os 4 primeiros blocos 4x4 estão contidos em um bloco 8x8, os outro 4 blocos 4x4 formam outro bloco 8x8, até que quatro blocos 8x8 formam um macrobloco 16x16. Assim, é facilitado o trabalho de agrupar os blocos em função da semelhança de seus vetores. Desse modo, a arquitetura desenvolvida recebe quatro vetores de movimento de blocos 4x4 e define de que modo é possível agrupar esses vetores.

No **decisor de partição**, primeiramente são comparados o primeiro vetor com o segundo e o terceiro com o quarto, para verificar a possibilidade de formar blocos 4x8. Ao mesmo tempo, é comparado o primeiro vetor com o terceiro, e o segundo vetor com o quarto, para verificar se é possível formar blocos 8x4. Caso o resultado de todas as comparações resultarem em possibilidade de agrupamento (por igualdade ou semelhança), significa que os quatro blocos serão agrupados em um bloco 8x8. Caso os 16 vetores recebidos resultem em quatro blocos 8x8, então o campo "Flag" passa a conter o valor 1 e os vetores dos quatro blocos 8x8 são

reaplicados à arquitetura para que seja definida sua forma de agrupamento. O processo é idêntico ao apresentado para o agrupamento de blocos 4x4. A diferença é que o resultado dos possíveis agrupamentos serão blocos de tamanho 8X16, 16x8 ou 16x16.

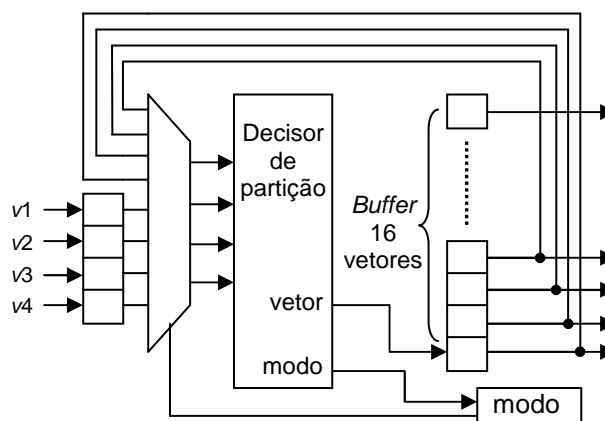
A Tabela 1 apresenta os códigos gerados pela arquitetura para sinalizar o tipo de agrupamento adotado. O campo “Flag” indica se foram usadas partições de bloco maiores que 8x8 (flag=1) ou partições com tamanho maior que 8x8 (flag=0). O campo “Código” refina a informação. Quando Flag=0, o Código indica se a **sub-partição** do bloco 8x8 será 4x4 (Código=00), 4x8 (Código=10), 8x4 (Código=01) ou 8x8 (Código=11). Com Flag=1, a informação do Código indica qual a **partição** utilizada sob o macrobloco 16x16, podendo ser 8x8 (Código=00), 8x16 (Código=10), 16x8 (Código=01) ou 16x16 (Código=11). A composição do “Flag” com os códigos gera um **modo** em 9 bits.

Tabela 1. Codificação do agrupamento de blocos

| Partição | Flag | Código |
|----------|------|--------|
| 4x4 | 0 | 00 |
| 8x4 | 0 | 01 |
| 4x8 | 0 | 10 |
| 8x8 | 0 | 11 |
| 8x8 | 1 | 00 |
| 16x8 | 1 | 01 |
| 8x16 | 1 | 10 |
| 16x16 | 1 | 11 |

A Figura 1 apresenta o diagrama em blocos da arquitetura. Os 4 vetores v_i são a entrada da arquitetura. O decisor de partição é composto por um uma série de comparadores e portas lógicas para definir a igualdade ou semelhança entre os vetores. O decisor gera os valores apresentados na Tabela 1. Esses valores são agrupados no registrador **modo**. Os vetores escolhidos vão para um buffer *ping-pong* de 16 posições a fim de serem armazenados. No caso da decisão resultar em “Flag=1”, então, os primeiros 4 vetores do buffer devem realimentar a arquitetura para que seja feita a decisão da partição de macrobloco. Esta nova decisão sobrescreve o modo e os vetores.

Figura 1. Diagrama em blocos do arquitetura desenvolvida



3. RESULTADOS E DISCUSSÃO

A arquitetura foi descrita em VHDL e sintetizada para um dispositivo FPGA, modelo EP2S15F484C3 da família Stratix II, utilizando o software Quartus II (ALTERA, 2009).

Com a síntese foi atingida uma frequência de 210,7 MHz. Isto significa que esta arquitetura possui uma taxa de processamento de aproximadamente 210 milhões de vetores por segundo. Foram utilizados 47 registradores e 57 ALUTs, o que representa menos de 1% da capacidade do dispositivo.

Considerando vídeos de alta resolução, como *Full* HD (resolução de 1920x1080 *pixels*), a arquitetura desenvolvida é capaz de decidir a partição de macrobloco para até 1619 quadros por segundo, taxa que ultrapassa facilmente a restrição de processamento em tempo real (30 quadros por segundo).

4. CONCLUSÕES

Este trabalho apresenta o projeto de uma arquitetura baseada em heurística para o modo de decisão da Estimação de Movimento do padrão H.264/AVC. Partindo dos 16 vetores 4x4 de um macrobloco, a arquitetura aqui apresentada decide o tipo de partição de macrobloco que deverá ser usada na Compensação de Movimento.

Como trabalho futuro, pretende-se testar a eficiência em termos de compressão de vídeo do algoritmo proposto (utilizando o software de referencia do padrão H.264/AVC – JM) (ITU-T, 2009). É previsto o desenvolvimento de novos algoritmos e a integração das arquiteturas geradas com o módulo de estimação de movimento VBSME (PORTO, 2009).

5. REFERÊNCIAS BIBLIOGRÁFICAS

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.

FÓRUM DO SISTEMA BRASILEIRO DE TV DIGITAL TERRESTRE. Disponível em: <<http://www.forumsbtvd.org.br>> Acesso em: Agosto 2009.

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

ALTERA CORPORATION: Altera - The Programmable Solutions Company. Disponível em: <<http://www.altera.com>> Acesso em: Agosto 2009.

Video Team Reference Software JM 16.0, ITU-T, 2009. Disponível em <<http://iphome.hhi.de/suehring/tml/>> Acesso em: Agosto 2009.

PORTO, Roger, "Hardware Design of the H.264/AVC Variable Block Size Motion Estimation for Real-Time 1080HD Video Encoding," *isvlsi*, pp.115-120, 2009 IEEE Computer Society Annual Symposium on VLSI, 2009.

AGOSTINI, Luciano, "Desenvolvimento de Arquiteturas de Alto Desempenho Dedicadas à Compressão de Vídeo Segundo o Padrão H.264/AVC", TESE DE DOUTORADO.