

DESENVOLVIMENTO DE UM COMPRESSOR DE QUADROS DE REFERÊNCIA UTILIZANDO CODIFICAÇÃO DIFERENCIAL PARA COMPRESSÃO DE VÍDEOS DIGITAIS

GUILHERME POVALA; LIVIA AMARAL, DIEISON SILVEIRA, BRUNO ZATT, MARCELO PORTO; LUCIANO AGOSTINI

Universidade Federal de Pelotas

{gpovala, lsdamaral, dssilveira, zatt, porto, agostini}@inf.ufpel.edu.br

1. INTRODUÇÃO

Com o crescimento na demanda por vídeos com melhor qualidade e com maiores resoluções, a codificação de vídeo vem sendo cada vez mais um processo fundamental para tornar vídeos digitais viáveis. Sem esse processo, os vídeos requereriam um volume de dados muito grande, e essa informação precisa ser armazenada e eventualmente transmitida, o que faria com que não fosse possível, por exemplo, transmitir vídeos pela internet de forma eficiente. Além disso, dispositivos móveis capazes de gravar e transmitir vídeos são cada vez mais comuns na vida das pessoas, como smartphones, câmeras digitais, tablets e outros.

O processo de codificação de vídeos é dado por uma variedade de ferramentas e técnicas, que foram e ainda são desenvolvidas, tornando a pesquisa nesta área muito ativa. E são com essas técnicas e ferramentas que os vídeos digitais são representados com um tamanho muito menor do que precisariam, a custo de uma pequena perda de qualidade. Os padrões de codificação de vídeos são definidos associando um conjunto dessas ferramentas e técnicas. O padrão H.264/AVC (ISO/IEC, 2012) é o mais utilizado no mercado atualmente, além de ter uma grande área de pesquisa no assunto. Porém, durante os últimos anos, o novo padrão de codificação de vídeo, o HEVC (*High Efficiency Video Coding*) (JCT-VC, 2013) foi e vem sendo desenvolvido, e neste ano, aprovado. Com o novo padrão de codificação de vídeos, o HEVC, é possível alcançar o dobro da taxa de compressão comparado com o H.264/AVC, mantendo a mesma qualidade (JCT-VC, 2013).

O processo de codificação de vídeos é realizado através de vários módulos, que utilizam uma grande quantidade de memória externa para realizar a codificação, especialmente a ME (*Motion Estimation*), que é responsável pela maior parcela de comunicação com a memória no codificador de vídeo. E isso se deve ao fato de que os quadros codificados precisam ser armazenados na memória, já que são utilizados como quadros de referência durante o processo de codificação da ME.

Na medida em que a comunicação com a memória impacta significativamente tanto com consumo de energia quanto por tempo de processamento, o codificador deve se preocupar com esse gargalo no sistema, que é a comunicação com a memória externa e as unidades de processamento.

Existem soluções para diminuir a quantidade de acessos à memória, (1) reuso de dados através de memórias caches, e (2) comprimindo os quadros de referência antes de serem armazenados na memória externa. A vantagem da segunda estratégia é que ela reduz tanto a quantidade de escritas quanto leituras

da memória, enquanto que a primeira estratégia, reduz apenas a quantidade de leituras. Porém, as duas estratégias podem ser utilizadas em conjunto.

Esse trabalho tem como objetivo propor um algoritmo para a redução de banda através da compressão de quadros de referência, com baixo custo computacional, de fácil implementação em hardware e sem inserir perdas no processo de codificação, denominado RFVLD (Reference Frame Variable-Length Differential Coder). Essa solução segue a ideia do trabalho apresentado em (SILVEIRA, 2013). O RFVLD é um codificador diferencial e utiliza codificação de Huffman para otimização dos códigos com foco nos padrões atuais de codificação de vídeo, como o H.264/AVC e o HEVC, os quais demandam uma grande quantidade de memória.

O algoritmo do RFVLD é aplicado sobre blocos de tamanho 12x12 e é capaz de reduzir mais de 38% do volume de dados armazenados na memória externa, assim, diminuindo a quantidade de operações de leitura e escrita.

2. METODOLOGIA

Nesse trabalho, a compressão dos quadros de referência é realizada sobre blocos 12x12 com sub-divisões de blocos 4x4. Com esses blocos de tamanho 4x4, são obtidos os resíduos (diferença entre cada amostra com uma amostra de referência) a partir do bloco 4x4 central. Foram avaliados quatro modos de calcular os resíduos.

Para cada modo, uma tabela estática de Huffman foi gerada a partir dos resíduos calculados. As sequências de vídeo utilizadas nesse trabalho foram codificadas com o software de referência HM 12.0 do padrão HEVC no modo lowdelay main, com QP 32. Essas sequências tem resolução 1080p, ou *full* HD (1920x1080 pixels) e todos os seus quadros foram utilizados. As sequências foram: BQTerrace, Blue Sky, Cactus, In to Tree, Pedestrian Area, Riverbed, Sun Flower e Tractor. Esses vídeos apresentam diferentes características de iluminação e movimentação, e são amplamente conhecidas pela comunidade científica da área.

Em codificadores de vídeo que atingem codificação em tempo real, por questões de otimizações, apenas amostras de luminância são utilizadas pela ME, e essas amostras são representadas por 8 bits, podendo assumir 256 diferentes valores. Como as tabelas geradas são diferenças entre amostras, essa diferença pode variar entre -255 e 255, assim, cada tabela é formada por 512 posições, uma para cada resíduo possível calculado entre amostras.

Na Figura 1, são apresentados os quatro modos de predição utilizados para a obtenção dos resíduos e para realização da codificação do quadro. No modo A, para cada amostra de um bloco 4x4 é calculado o resíduo a partir da amostra na mesma posição do bloco referência (bloco central mostrado na Figura 1), conforme pode ser visto na Figura 1A. Além disso, o modo A aplica a compressão do bloco referência a partir de oito tabelas estáticas, de acordo com (SILVEIRA, 2013).

O modo B, opera da mesma forma que o modo A, onde cada amostra de um bloco 4x4 é subtraída da amostra de mesma posição do bloco de referência, de acordo com a Figura 1B, porém, para o bloco referência nenhuma forma de compressão é aplicada. Assim, apenas a tabela utilizada nos resíduos é necessária. Como cada uma das oito tabelas utilizadas em (SILVEIRA, 2013) tem 256 posições, diminuimos 80% da quantidade de espaço necessário para armazenar as 8 tabelas mais a tabela de resíduos, que é o caso do modo A. Porém, a taxa de compressão é menor, já que o bloco referência não é

comprimido. No modo C, cada amostra de um bloco 4x4 é subtraída da média do bloco referência, como na Figura 1C, e esse, armazenado sem compressão.

Já no modo D, cada amostra de um bloco 4x4 é subtraída da amostra correspondente a borda adjacente do bloco referência, conforme pode ser visto na Figura 1D. Nesse modo, foi obtida a melhor taxa de compressão, inclusive, superior a taxa de compressão obtida em (SILVEIRA, 2013). Isso é explicado por que a diferença é obtida de amostras muito próximas, fazendo com que os resíduos sejam, em sua maioria, próximos de zero, de acordo com a Figura 2. Além do aumento na taxa de compressão, houve a diminuição de 75% no espaço necessário para armazenar as tabelas comparando com (SILVEIRA, 2013).

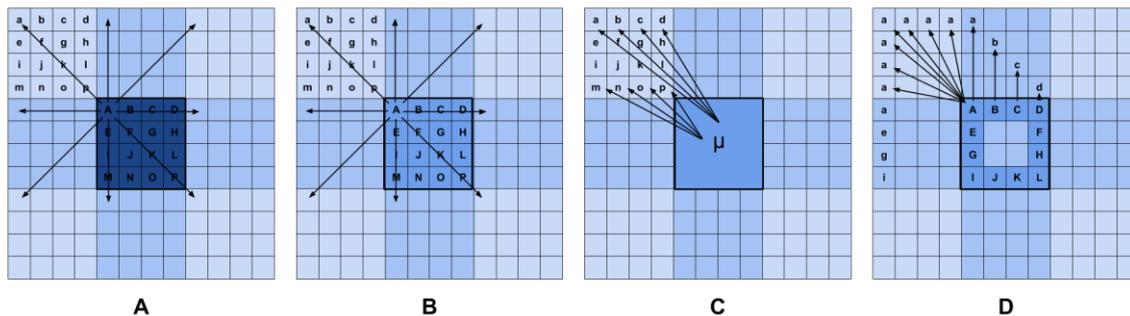


Figura 1 – Modos de Predição Utilizados

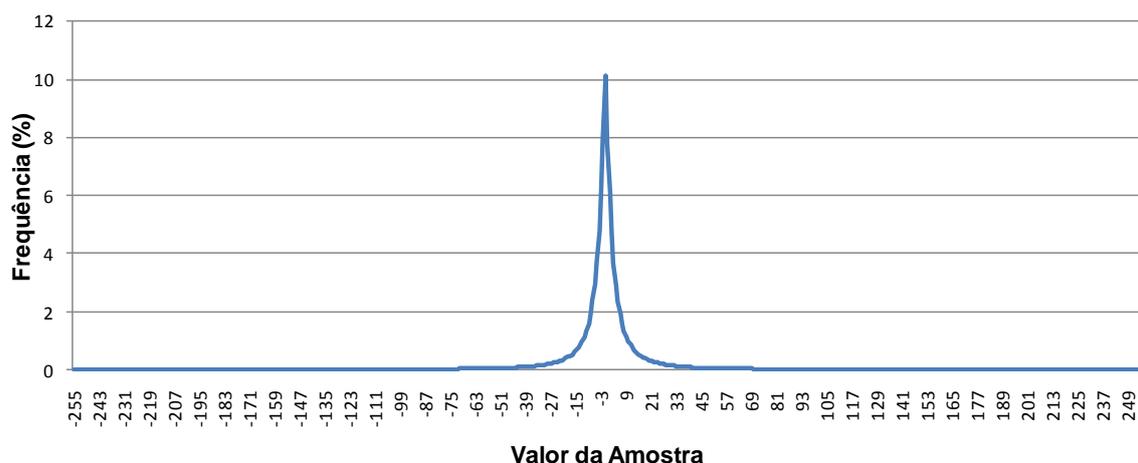


Figura 2 – Frequência dos Resíduos no Modo D

3. RESULTADOS E DISCUSSÃO

A Tabela 1 apresenta os resultados de taxa de compressão obtidos pela técnica desenvolvida. Os resultados foram obtidos utilizando oito vídeos 1080p, com diferentes aspectos de movimento e iluminação. Além disso, quatro desses vídeos não participaram na criação da tabela estática de resíduos, isso foi feito de forma intencional, a fim de provar que essa solução independe do vídeo utilizado. A taxa de compressão obtida também indica a porcentagem de diminuição na largura de banda de memória externa.

Como pode ser visto, uma variação ocorre na taxa de compressão entre os vídeos. Isso ocorre pelo fato de que diferentes vídeos apresentam diferentes comportamentos, um deles é a quantidade de áreas homogêneas. Dependendo dessas áreas, uma maior ou menor taxa de compressão é obtida, já que é calculado o resíduo entre amostras, portanto, como pode ser visto na Figura 2,

quanto menor o resíduo, maior é sua ocorrência e consecutivamente, menor o seu código de representação.

A taxa média de compressão ficou em 38,16%, com um desvio padrão de 6,29%, indicando que essa solução atinge boas taxas de compressão independente do vídeo avaliado. Isso significa que o tempo necessário para a ME realizar os acessos à memória pode ser reduzido em 38,16%.

Tabela 1 – Taxas de Compressão Aplicando Diferentes Métodos de Predição

Vídeo	Taxa de Compressão	Vídeo	Taxa de Compressão
Blue Sky	31,06%	Kimono	37,26%
In to Tree	42,23%	Rolling Tomatoes	46,85%
Pedestrian Area	38,85%	Station2	36,44%
Sun Flower	32,85%	Tennis	39,76%
Média de Compressão		38,16%	

Comparando essa solução com o trabalho apresentado em (SILVEIRA, 2013), temos um ganho de 20,61% em na taxa de compressão obtida por esse trabalho, além de diminuir em 75% a área necessária para armazenar as tabelas de códigos de Huffman.

4. CONCLUSÕES

Esse trabalho apresentou uma solução para reduzir a largura de banda de memória em codificadores de vídeos digitais, utilizando para isso, a compressão de quadros de referência. O método apresentado não insere perdas de qualidade no processo de codificação, e para isso, realiza uma codificação diferencial inter-blocos, seguindo uma estratégia de proximidade das amostras dos blocos através das bordas do bloco de referência.

Para a codificação, uma tabela estática de Huffman de 512 posições é utilizada, e é realizada sobre blocos de 12x12 amostras, com blocos internos de 4x4, atingindo uma taxa de compressão média de 38,16% e, consecutivamente, a mesma redução em largura de banda de memória.

Como trabalho futuro, será estudada uma forma de comprimir o bloco de referência, aumentando ainda mais a taxa de compressão. Além disso, a solução proposta será implementada em VHDL e sintetizada para um dispositivo FPGA e ASIC para obtenção de resultados de consumo de área e energia.

5. REFERÊNCIAS BIBLIOGRÁFICAS

ISO/IEC. **International Organization for Standardization. ISO/IEC 14496-10 mpeg-4 part 10 - coding of audio-visual objects - part 10: Advanced video coding.** Technical Report, 2012.

JCT-VC. **Joint Collaborative Team on Video Coding.** Acessado em: 02 out. 2013. Online. Disponível em <http://phenix.int-evry.fr/jct/index.php>.

SILVEIRA, D.; PORTO, M.; AGOSTINI, L. A lossless approach for external memory bandwidth reduction in video coding systems and its VLSI architecture. **IEEE ICME 2013**, San Jose, USA, 2013.