

PADRÃO DE BUSCA ASTROIDE PARA ESTIMAÇÃO DE MOVIMENTO NO CODIFICADOR DE VÍDEO HEVC

PAULO GONÇALVES; GUILHERME CORREA; MARCELO PORTO

Universidade Federal de Pelotas – Video Technology Research Group (ViTech)
{phrgoncalves, gcorrea, porto}@inf.ufpel.edu.br

1. INTRODUÇÃO

Frente a crescente demanda por avanços nas tecnologias de vídeos digitais, principalmente devido as diversas aplicações de vídeos de alta resolução em dispositivos portáteis, como smartphones e tablets, e serviços de streaming, o consumo por esta mídia se tornou uma parte rotineira no cotidiano de grande parte das pessoas. Contudo, é inviável transmitir e armazenar vídeos digitais quando esses estão em seu formato original, e nesse contexto o uso de compressores de vídeos digitais torna-se praticamente imprescindível.

A compressão de vídeo é realizada explorando redundâncias temporais e espaciais, podendo assim atingir altas taxas de compressão (SULLIVAN, 2012). Atualmente, um dos padrões de codificação de vídeo mais explorados é o *High Efficiency Video Coding* (HEVC) (JCT-VC, 2013), pois apresenta uma eficiência de compressão 40-50% maior que seu antecessor, o padrão H.264/AVC. Em contrapartida, esse aumento de eficiência veio ao custo de um aumento significativo de complexidade, que dependendo da configuração utilizada no codificador pode ser de 500% (CORRÊA, 2012).

Uma das etapas presentes nos codificadores de vídeo atuais, é a **Estimação de Movimento** (ME, do inglês *Motion Estimation*), a qual consiste em reduzir as redundâncias temporais, ou seja, diminuir a quantidade de dados repetidos nos diferentes quadros de um vídeo. Esta é uma das etapas que mais demanda tempo em um codificador de vídeo, e no HEVC, um dos motivos para este aumento de complexidade é a utilização de 24 diferentes tamanhos de blocos na realização da ME.

O processo de particionamento e subparticionamento das unidades de codificação segue uma hierarquia, onde cada quadro de um vídeo é dividido em *Coding Tree Units* (CTUs), que são unidades de dimensão 64×64 (pixels). Estas, por sua vez, são divididas em *Coding Units* (CUs), cujos tamanhos podem variar entre 64×64, 32×32, 16×16 e 8×8. Estas CUs também são futuramente divididas em *Prediction Units* (PUs), que diferente das outras partições, podem apresentar formatos assimétricos. A etapa de ME é responsável por buscar o bloco mais similar ao bloco que está sendo codificado no quadro atual em um quadro de referência. Essa similaridade pode ser calculada através de diferentes métricas, sendo a Soma das Diferenças Absolutas (SAD) mais comumente utilizada. Uma vez encontrado, o bloco com maior similaridade tem suas coordenadas armazenadas em um Vetor de Movimento (MV, do inglês *Motion Vector*).

Existem diversos algoritmos para a realização da etapa da ME, dentre eles o algoritmo **Test Zone Search** (TZS), o qual está implementado no software de referência do HEVC, o *HEVC Model* (HM). O algoritmo do TZS consiste em, basicamente, quatro etapas: A **Predição de Vetores de Movimento**, que busca apontar a região mais próxima ao possível melhor resultado. Esta região é chamada de Área de Busca; a etapa de **Busca Inicial**, que consiste em uma busca em expansão na região apontada pela etapa anterior até atingir o ponto de

parada. Neste momento são comparadas as distâncias do melhor bloco candidato até o momento e do bloco encontrado após a etapa de Predição, se essa distância for maior que cinco então executa-se a etapa de **Busca Raster** (que consiste em uma busca semelhante à busca exaustiva, mas de forma subamostrada, por toda a área de busca), e a seguir, o **Refinamento**. Caso a distância for menor que cinco, então executa-se somente a etapa de Refinamento, a qual é semelhante a Busca Inicial, também realizando uma busca em expansão, porém a faz de forma iterativa, ou seja, a cada iteração, se um novo bloco é tido como melhor candidato, então na próxima iteração, este bloco é escolhido como ponto inicial.

Tendo conhecimento do funcionamento e das características do algoritmo do TZS, este trabalho propõe um novo padrão de busca astroide para substituir a etapa de Busca Raster no TZS.

2. METODOLOGIA

Para extrair resultados sobre as características do algoritmo TZS, duas análises foram realizadas. Nestas análises, foi utilizado o software de referência do HEVC, o HM na versão 16.14. Ao todo, oito sequências de vídeo foram analisadas, todas com resoluções de alta e ultra definição. Também foram codificados todos os quadros utilizando os Parâmetros de Quantização (QP): 22, 27, 32 e 37, conforme recomendado nas Condições Comuns de Teste (CTC) (SHARMAN, 2018). Primeiramente, foi realizado uma análise sobre o tempo de execução de cada etapa do algoritmo TZS. A Figura 1 mostra os resultados obtidos para as oito sequências de vídeo citadas anteriormente, além do resultado médio. Para a sequência *Beauty*, a etapa de Busca Raster é responsável por 84,85% do tempo de execução do algoritmo TZS. Isto deve-se principalmente as características do vídeo, onde se tem o rosto de uma mulher muito próxima a câmera. Desta forma, os vetores de movimento na sua maioria, terão valores maiores que 5, e, portanto, a Busca Raster será habilitada diversas vezes. Análogo a isto, a sequência *HoneyBee* mostrou que a Busca Raster é responsável por apenas 15,45% do tempo de execução. Isto deve-se também a característica do vídeo, onde têm-se um campo florido com apenas uma abelha movimentando-se entre as flores. Desta forma, devido a pouca movimentação na cena, a Busca Raster quase nunca é habilitada. Em média, a Busca Raster mostrou-se responsável por **67,85%** de todo o tempo de execução TZS.

A etapa de Busca Raster é importante durante a execução do TZS, pois consegue alcançar limites da área de busca que as outras etapas na maioria das

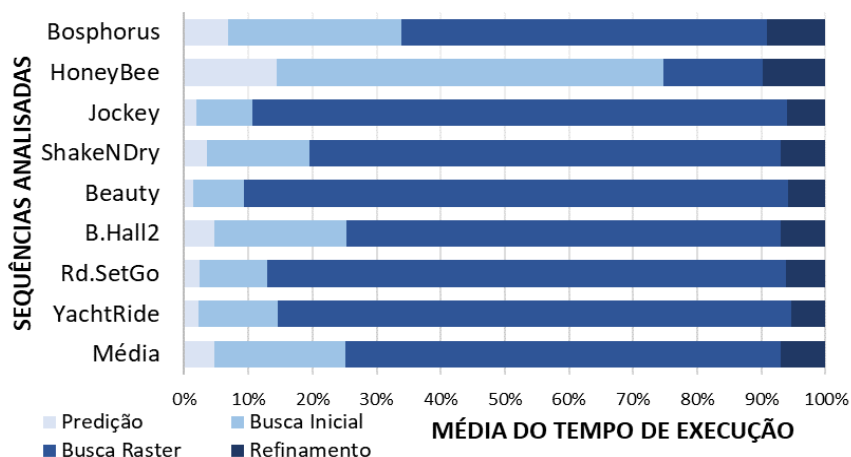


Figura 1: Distribuição do tempo de processamento do TZS em suas quatro etapas

vezes não alcançam. Desta forma, somente eliminar esta etapa não seria uma boa estratégia para redução do esforço computacional. Sendo assim, foi realizada uma análise do comportamento da Busca Raster, visando identificar a distribuição das melhores correspondências de blocos retornadas a partir desta etapa.

A análise da distribuição das melhores correspondências de blocos consiste em coletar, ao longo da execução da etapa de Busca Raster, a posição da melhor correspondência dentro da área de busca. O tamanho da área de busca define o número de comparações que será realizado na etapa de Busca Raster, por exemplo, em uma área de busca de tamanho 512×512 posições, considerando uma subamostragem de 5 posições, como implementado no TZS, aproximadamente 10.500 posições são testadas durante uma execução desta etapa. Os resultados foram traçados em mapas de calor, e são apresentados na Figura 2, onde as cores mais quentes indicam a área de maior ocorrência de melhor correspondência de bloco e as cores mais frias apresentam a área que foi raramente ou nunca escolhida. A Figura 2(a) apresenta o mapa de calor para a sequência de vídeo *HoneyBee*, que possui a distribuição mais incomum entre todas as sequências analisadas, onde é possível observar elevações por toda a área de busca. Isto deve-se ao fato desta sequência realizar poucas vezes a etapa de Busca Raster, assim como já mencionado anteriormente, e, portanto, o gráfico se torna mais suscetível a mudanças. A Figura 2(b) mostra a média obtida para todas as 8 sequências analisadas. Neste caso, é possível perceber uma grande concentração na área central e nos eixos, e possibilita a visualização de um padrão em formato diamante.

Esta análise mostrou que o esforço computacional para calcular grande parte dos pontos dentro da Busca Raster é desperdiçado, sendo que a maioria dos pontos comparados fora do centro e dos eixos da área de busca dificilmente são escolhidos como melhores, sendo assim, é perceptível que esses pontos podem ser removidos, com baixas perdas em eficiência de codificação. Para explorar as características observadas na análise anterior de forma eficiente, este trabalho propõe um novo padrão de busca astroide. Conforme ilustrado na Figura 2(c), um Astroide (demarcado em azul) é uma hipocicloide com quatro vértices, traçada por um ponto na circunferência de um círculo C_1 (demarcado em verde), enquanto este rola sem escorregar no interior de um segundo círculo fixo maior C_2 (traçado em verde). Este padrão é muito similar ao observado na média apresentado na Figura 2(b), e foi desenvolvido para explorar somente 25% da área de busca original, que corresponde ao traçado em vermelho na Figura 2(c). Contudo, mesmo realizando somente 25% do total de comparações, o padrão astroide é capaz de cobrir em média 72,4% de todas as melhores correspondências de blocos encontradas na etapa de Busca Raster, considerando as mesmas sequências de vídeo analisadas anteriormente.

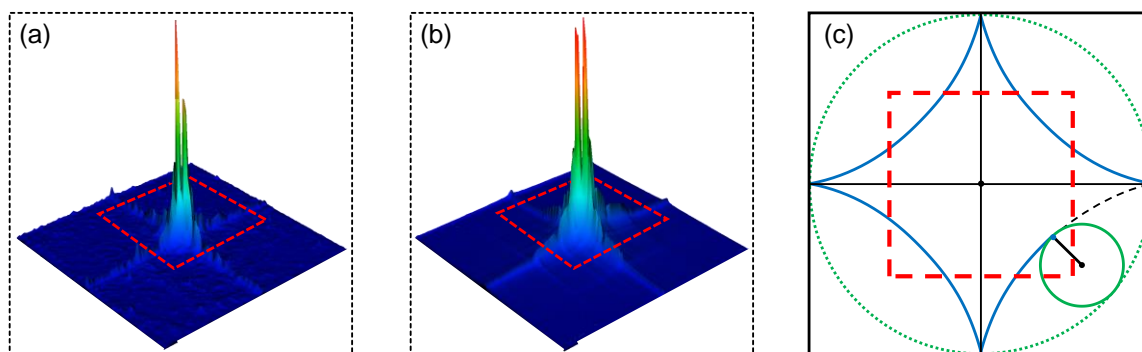


Figura 2: (a) Distribuição de correspondência de bloco na área de busca 512×512 para a sequência *HoneyBee*; (b) Distribuição média para todas as sequências analisadas; (c) Padrão de Busca Astroide Proposto

3. RESULTADOS E DISCUSSÃO

Para validar a eficiência do esquema proposto, foram utilizadas dezessete sequências de vídeo de ultra e alta definição, das classes A1, A2, A e B recomendadas na CTC (SHARMAN, 2018). As mesmas condições mostradas na seção 2 foram respeitadas. É importante enfatizar que todas as sequências são diferentes das utilizadas para as análises apresentadas neste trabalho, e diferem em taxa de quadros, profundidade de bits, resolução espacial e conteúdo de movimento/textura. A implementação do HM modificado com a solução proposta foi comparada ao original para permitir comparações em termos de redução de tempo total de codificação e redução de tempo somente para o algoritmo TZS. A eficiência de compressão foi calculada em termos de BD-rate (BJØNTEGAARD, 2011), que mostra a variação da taxa de bits do HM modificado para manter a mesma qualidade visual do HM original.

Os resultados do padrão proposto são mostrados na Tabela I. O esquema proposto atinge redução de tempo em média de **46,27%** para o TZS e **12,89%** para todo o processo de codificação, com um BD-rate de **-0,2162%** em relação ao software de referência, o que significa que a solução proposta consegue diminuir a taxa de bits, mantendo a mesma qualidade visual da versão original.

Tabela 1: Comparações entre o Padrão Proposto e o TZS original

Classe	Resolução	BD-rate (%)	Redução de Tempo TZS (%)	Redução de Tempo Total (%)
A1	3840x2160 (Ultra HD)	-0,3020	54,82%	19,11%
A2	3840x2160 (Ultra HD)	-0,0047	45,68%	12,40%
A	2560x1600 (WQXGA)	-0,3162	42,97%	10,33%
B	1920x1080 (Full HD)	-0,2242	41,63%	09,72%
Média		-0,2162	46,27%	12,89%

4. CONCLUSÕES

Este trabalho apresentou um novo padrão de busca astroide para substituir a etapa de Busca Raster no algoritmo *Test Zone Search*, a fim de reduzir o esforço computacional presente na etapa de Estimação de Movimento em codificadores de vídeo atuais. Com base nos resultados apresentados, é possível afirmar que o esquema obteve um ganho em redução de tempo considerável quando comparado a versão original, além de uma redução na taxa de bits para a mesma qualidade visual.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- SULLIVAN et al. **Overview of the High Efficiency Video Coding (HEVC) Standard**; IEEE Transactions on Circuits and Systems for Video Technology. [S.l.:s.n], v. 22 n.12, p.1649-1668, 2012.
- JCT-VC. **High Efficiency Video Coding text specification draft 10**, doc. JCTVC-L1003, Genebra, Suíça, 2013.
- CORRÊA, G.; et al. **Performance and Computational Complexity Assessment of High Efficiency Video Encoders**. IEEE Transactions on Circuits and Systems for Video Technology, v. 22, n. 12, p. 1899-1909, 2012.
- SHARMAN, K. K. S. **Common Test Conditions for HM video coding experiments**. JCT-VC Document JCTVC-AC1100.: [s.n.], 2018.
- BJØNTEGAARD, G. **Calculation of Average PSNR Differences between RD-curves**. ITU Documents, Austin 2011. Acessado em 05 julho de 2017, Online. Disponível em: wftp3.itu.int/av-arch/video-site/0104_Aus/VCEG-M33.doc.