

## ANÁLISE DA DISTRIBUIÇÃO DOS MODOS DE PREDIÇÃO INTRA EM CODIFICADORES DE VÍDEO SEGUNDO O PADRÃO HEVC

MARILIA PIT; GUILHERME CORREA;  
DANIEL PALOMINO; MARCELO PORTO

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

### 1. INTRODUÇÃO

A codificação de vídeos tem como objetivo a diminuição da quantidade de dados considerados redundantes na representação computacional das informações do vídeo. É considerado redundante o dado que não contribui com novas informações relevantes para a representação do vídeo. Existem três tipos diferentes de redundâncias exploradas na compressão de vídeos: espacial, temporal e entrópica. A redundância espacial, especialmente relevante para este trabalho, é também chamada de “redundância intra-quadros” e advém da correlação existente entre os pixels espacialmente distribuídos em um frame, os quais, se próximos, tendem a serem muito semelhantes. Desta forma, o módulo de codificação intra-quadro é responsável por reduzir a redundância espacial do frame em processamento (AGOSTINI, 2007).

O *High Efficiency Video Coding* (HEVC) (JCT-VC, 2013), padrão desenvolvido pelo *Joint Collaborative Team on Video Coding* (JCT-VS) e atual estado-de-arte da compressão de vídeos, define que cada frame seja particionado em *Coding Tree Units* (CTUs), as quais são particionadas em *Coding Units* (CUs), que dividem regiões codificadas com predição intra-quadro e com predição inter-quadros. Cada CU, quando intra-codificada, consistirá em uma *Prediction Unit* (PU)  $2N \times 2N$  ou quatro PUs  $N \times N$ , tamanho subordinado à sua profundidade na árvore de particionamento, onde  $N$  representa metade da CU. A predição intra-quadro do HEVC suporta PUs de  $4 \times 4$  a  $64 \times 64$  pixels e cada PU pode empregar um diferente modo de predição. São 35 modos possíveis – 33 angulares, para contemplar diferentes estruturas direcionais tipicamente presentes em vídeos e imagens, um DC, apropriado para regiões de textura homogênea, e um Planar, apropriado para regiões suaves e de mudanças graduais.

Para determinar o melhor tamanho e o melhor modo de predição de cada PU, o HEVC segue as seguintes etapas: a *Rough Mode Decision* (RMD), a inclusão do *Most Probable Modes* (MPM) e o cálculo do *Rate-Distortion Optimization* (RDO). A RMD é aplicada aos 35 modos para gerar uma lista com um subconjunto dos  $N$  modos mais prováveis para serem aplicados ao cálculo do RDO, os quais são selecionados embasados no custo *Hadamard* de cada um dos 35 modos e na estimativa de consumo de bits, de acordo com a equação (1).

$$C_i = SATDMode_i + \lambda * BitsMode_i \quad (1)$$

onde *SATDMode<sub>i</sub>* é a soma das transformadas Hadamard do modo de predição  $i$ ,  $\lambda$  é o multiplicador de Lagrange e *BitsMode<sub>i</sub>* é a estimativa de custo de bits do modo de predição  $i$ .

Em razão da redundância espacial, é bastante provável que o modo ótimo para a codificação da PU atual seja o mesmo das PUs vizinhas já codificadas. Assim, os modos da vizinha à esquerda e da vizinha superior (MPM) são

incorporados à lista de candidatos para o cálculo do RDO. Além disso, um terceiro modo é acrescentado ao subconjunto de modos na lista MPM, sendo este Planar, DC ou Angular (26), nesta ordem. Caso haja modos duplicados, a substituição também é feita nesta ordem e se a PU vizinha superior e a PU vizinha à esquerda foram codificadas pelo mesmo modo, este modo e os dois modos angulares mais próximos são utilizados para a construção do subconjunto.

Por fim, o cálculo do RDO (SULLIVAN et al., 1998) é realizado sobre os modos de predição presentes na lista de candidatos. Esta lista contém 8 modos para PUs  $4 \times 4$  e  $8 \times 8$ , a fim de permitir uma busca mais refinada em blocos pequenos, e 3 modos para PUs  $16 \times 16$ ,  $32 \times 32$  e  $64 \times 64$ . O modo de predição com o menor custo RD é o selecionado como modo ótimo para a efetiva codificação da PU.

A decisão de modo é responsável por mais da metade da complexidade computacional da codificação de um quadro intra no *HEVC Test Model* (SILVA et al., 2012). Desta forma, o objetivo desta pesquisa é o desenvolvimento de um novo algoritmo de decisão de modo intra, a fim de reduzir a complexidade desta etapa da codificação. Especificamente neste resumo, são apresentados os resultados preliminares da análise da distribuição dos modos de predição entre as PUs durante a codificação de quadros intra.

## 2. METODOLOGIA

O desenvolvimento algorítmico deu-se sobre o código fonte do *HEVC Test Model* (versão 16.16), o software de referência desenvolvido em conjunto pelo ITU-T *Video Coding Experts Group* e pelo ISO/IEC *Moving Picture Experts Group*.

O ponto de partida foi o estudo deste software, a fim de identificar e modificar as implementações responsáveis pela decisão de modo intra para possibilitar a extração de dados. No HEVC, as CUs de uma CTU são percorridas e codificadas de forma não linear, através da técnica de *Z-Scan*. Assim, no momento da codificação de uma PU, podem haver mais vizinhas já codificadas cujos modos podem ser relevantes para a construção da lista de candidatos, não se limitando apenas às duas consideradas pelo MPM. São estas as PUs vizinhas à esquerda-inferior, esquerda-superior e direita-superior. Foram então inseridas rotinas para coletar e comparar os modos da PU atual e de suas cinco vizinhas.

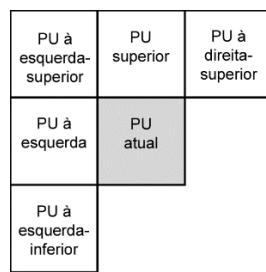


Figura 1: Disposição da PU atual e suas PUs vizinhas.

Como o custo estimado pela RMD é mais rápido de ser calculado, porém menos preciso que o do RDO, o subconjunto de modos na lista de candidatos pode não conter o modo ótimo para a predição da PU. Desta forma, foram descartadas as etapas de RMD e de MPM, sem preocupações com a eficiência computacional, com a finalidade de que todos os 35 modos fossem avaliados no cálculo do RDO.

Para analisar a relação entre o modo escolhido para a PU atual e os modos escolhidos nas suas cinco vizinhas foram codificados 50 quadros de duas

diferentes sequências de testes, *BasketballDrive* e *Cactus*, ambas com resolução  $1920 \times 1080$ , sob quatro parâmetros de quantização distintos, 22, 27, 32 e 37, na configuração *Intra Main*. Esta configuração garante que todas as CUs de todos os quadros do vídeo sejam codificadas utilizando somente a predição intra-quadro.

### 3. RESULTADOS E DISCUSSÃO

Conforme mencionado, os resultados e discussões apresentados nesta seção consideram duas sequências de teste: *BasketballDrive* e *Cactus*. As Figuras 2-4 apresentam uma compilação e análise dos resultados obtidos para os dois vídeos, sempre focando na relação entre a PU atual e as suas cinco PUs vizinhas.

A Figura 2 evidencia a ocorrência do modo ótimo da PU atual entre suas PUs vizinhas, a qual cresce à medida que o QP também cresce, devido à perda de informações que resulta numa menor precisão de codificação. Nas PUs de  $4 \times 4$  pixels, eficientes na codificação de regiões ricas em detalhes, a taxa de ocorrência do modo ótimo da PU atual entre suas PUs vizinhas supera 80% dos casos em ambos os vídeos no QP 37.

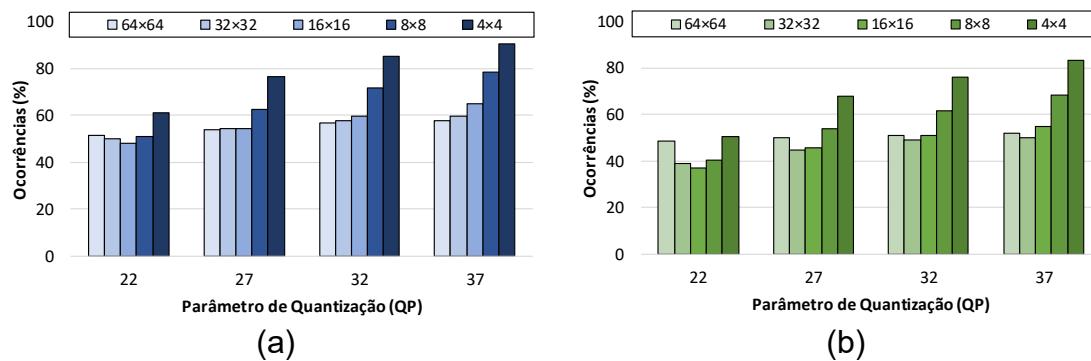


Figura 2: Ocorrência do modo de predição da PU atual nas PUs vizinhas para os vídeos (a) *BasketballDrive* e (b) *Cactus*.

De maneira análoga, a Figura 3 demonstra a relação do modo intra mais frequente entre as PUs vizinhas e o modo ótimo da PU atual. A ausência de PUs de  $64 \times 64$  pixels na figura deve-se ao fato de que estas não possuem vizinhas à esquerda-inferior no momento da codificação, já que no HEVC as CTUs são codificadas através da técnica de *Raster-scan*. Se exatamente cinco PUs vizinhas foram codificadas por um mesmo modo, as chances de que a PU atual seja codificada por este modo também supera os 80% em ambos os vídeos no QP 37.

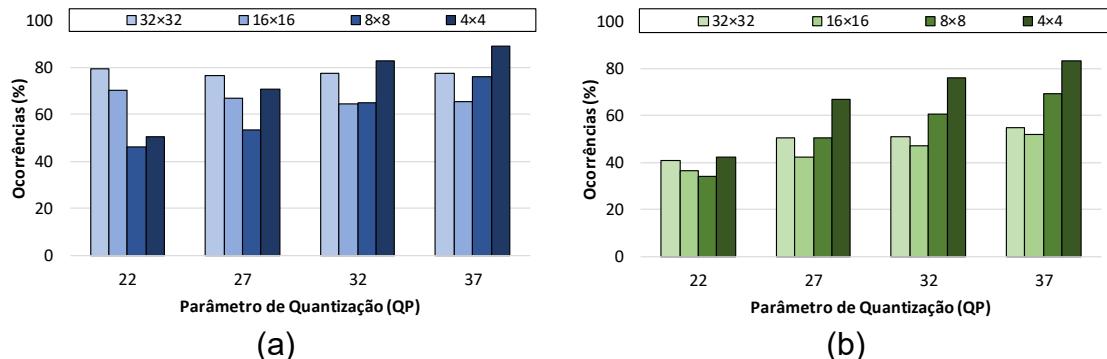


Figura 3: Ocorrência do mesmo modo entre as cinco PUs vizinhas coincide com o modo de predição da PU atual para os vídeos (a) *BasketballDrive* e (b) *Cactus*.

A probabilidade de que o modo intra majoritário entre as PUs vizinhas seja o modo ótimo para a PU atual decresce à medida que este modo torna-se menos frequente entre as PUs vizinhas. A Figura 4 demonstra a ocorrência do modo ótimo para a PU atual quando este mesmo modo aparece entre exatamente três PUs vizinhas.

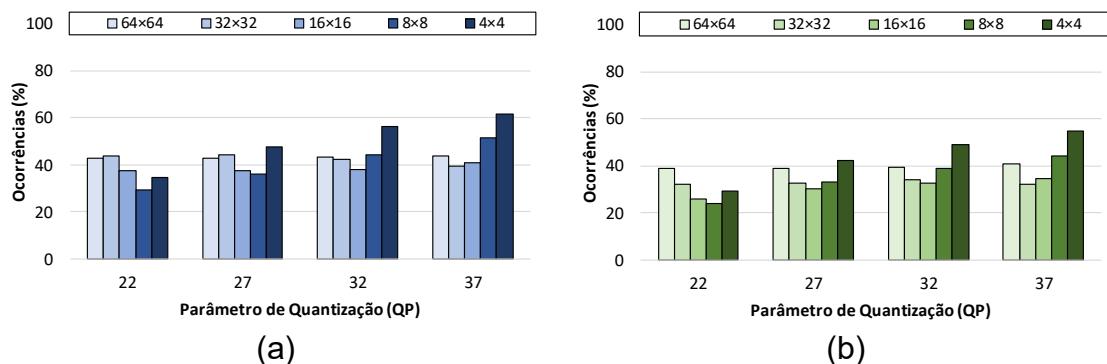


Figura 4: Ocorrência do mesmo modo entre as três PUs vizinhas coincide com o modo de predição da PU atual para os vídeos (a) *BasketballDrive* e (b) *Cactus*.

#### 4. CONCLUSÕES

Este trabalho apresentou uma análise da distribuição dos modos de predição intra-quadro no padrão de codificação de vídeo HEVC. A análise mostra que há uma grande probabilidade de acerto ao limitar a procura pelo modo ótimo da PU atual apenas entre os modos escolhidos nas PUs vizinhas. Entretanto, a análise também revela que em alguns casos o modo escolhido na PU atual não ocorre nas vizinhas. Em outros casos, mesmo que um modo ocorra com frequência entre as PUs vizinhas, ele acaba por não ser o modo ótimo escolhido para a PU atual. Assim, o próximo passo da pesquisa será analisar a relação dos custos RMD entre o modo da PU atual e os modos das PUs vizinhas.

A partir das análises e dos dados obtidos, a continuação desta pesquisa consiste em propor uma abordagem de redução de complexidade nas decisões da predição intra-quadro HEVC baseada na utilização de árvores de decisão construídas através de um processo de aprendizado de máquina.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

- AGOSTINI, L.V. **Desenvolvimento de Arquiteturas de Alto Desempenho Dedicadas à Compressão de Vídeo Segundo o Padrão H.264/AVC**. 2007. Programa de Pós-Graduação em Computação, Universidade Federal do Rio Grande do Sul. 2007.
- JCT-VC. High Efficiency Video Coding text specification draft 10, doc. JCTVC-L1003, Genebra, Suíça, 2013.
- DA SILVA, T. L.; AGOSTINI, L. V.; DA SILVA CRUZ, L. A. Fast HEVC intra prediction mode decision based on EDGE direction information. In: Proceedings of 20th European Signal Processing Conference (EUSIPCO), 2012, pp. 1214-1218.
- SULLIVAN, G. J.; WIEGAND, T. Rate-Distortion Optimization for Video Compression, **IEEE Signal Processing Magazine**, v15, n.6, p.74-90, 1998.
- CORRÊA, G.; ASSUNCAO, P. A.; AGOSTINI, L. V.; DA SILVA CRUZ, L. A. Fast HEVC encoding decisions using data mining. **IEEE Transactions on Circuits and Systems for Video Technology**, v25, n.4 p.660-673, 2015.