

ANÁLISE DE CORRELAÇÃO DAS ESTRUTURAS DE PARTICIONAMENTO NA TRANSCODIFICAÇÃO DE VÍDEOS HEVC PARA AV1

ALEX MACHADO BORGES; MARCELO PORTO; BRUNO ZATT;
GUILHERME CORRÊA

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

1. INTRODUÇÃO

Com o advento de novas tecnologias para captura e reprodução de vídeos, evoluem também as tecnologias para armazenamento e transmissão desses vídeos. Para tarefas de armazenamento são utilizados os codificadores de vídeos, que visam manter uma alta qualidade do vídeo em relação à captura original ao mesmo tempo que intenta na maior compressão possível. Atualmente a indústria e a academia têm utilizado os codificadores H.264/AVC (WIEGANG et. al; 2003) e H.265/HEVC (SULIVAN et. al.; 2012) como referência, já que atingem uma alta taxa de compressão com baixa perda visual na qualidade da imagem.

Todavia, há um custo elevado na utilização comercial desses codificadores de vídeo devido aos *royalties* cobrados pelos órgãos de padronização. Há outros codificadores no mercado, alguns até sem custos para a sua utilização, porém não tão eficazes como o HEVC. Cientes disso, algumas empresas de tecnologia resolveram unir forças e desenvolver um codificador de vídeo que fosse livre de *royalties* e com uma eficácia superior à atingida pelo HEVC. Surgiu a *Alliance for Open Media* e com isso iniciaram o desenvolvimento de seu primeiro codificador de vídeo: o *AMedia Video 1* (CHEN et. al.; 2018), apelidado de AV1.

Com o surgimento de um codificador de vídeo concorrente ao HEVC, que envolve no barateamento de custos para compressão e armazenamento desses vídeos, é do interesse de empresas que possuem seus vídeos codificados no padrão HEVC que estes sejam transformados para o novo codificador AV1. Esse procedimento é denominado transcodificação. A técnica tradicional de transcodificação, conhecida como tandem, consiste na realização de uma decodificação e uma codificação em sequência, como mostra a Fig. 1. Dado um vídeo codificado de acordo com o padrão A, por exemplo o HEVC, decodifica-se ele para obter o vídeo decodificado. Após, codifica-se o vídeo decodificado em um codificador que segue o padrão B, por exemplo o AV1.



Figura 1. Transcodificador A para B em tandem.

Apesar da simplicidade e eficiência da técnica tandem, ela exige um longo período de processamento do vídeo pelo novo codificador. Considerando o elevado tempo necessário para codificar um vídeo AV1, de aproximadamente 400 vezes mais que o HEVC (AKYAZI, 2018) (GROIS, 2016), a transcodificação tandem não é a mais recomendada. Porém é possível acelerar o processo fazendo uso de decisões já realizadas previamente no codificador original, desde que os codificadores de vídeo possuam similaridades de técnicas. Tanto o HEVC como o AV1 seguem a estrutura de um codificador genérico moderno com estruturas de particionamento de quadros bastante similares.

O HEVC divide o quadro em regiões de 64x64 píxeis, denominadas *Coding Tree Unit* (CTU), as quais podem ser recursivamente subdivididas em partes menores em uma estrutura de árvore quadrática, e assume quatro tamanhos válidos: 64x64, 32x32, 16x16 e 8x8. No AV1, o maior bloco possível é o superbloco, com 128x128 píxeis, que também pode ser recursivamente subdividido, a um total de 22 tamanhos de blocos diferentes. Na Fig. 2 é possível observar um exemplo das estruturas existentes em ambos os codificadores.

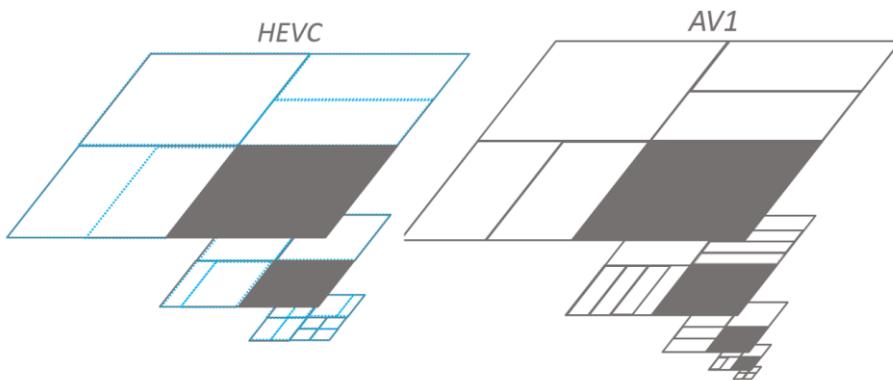


Figura 2. Exemplo de divisões de CTU (esquerda) e de superbloco (direita).

2. METODOLOGIA

Esta pesquisa intenta em analisar a possibilidade de reaproveitar as estruturas de particionamento decididas durante a codificação HEVC para limitar a estrutura de blocos do AV1. Isso é, dado uma certa profundidade de árvore quadrática de particionamento observada na codificação HEVC, pretende-se observar a viabilidade de não permitir que o AV1 explore todas as possibilidades.

Antes do desenvolvimento de uma solução, portanto, é preciso primeiramente determinar se há uma correlação real entre as árvores de particionamento dos dois codificadores em questão. De forma a possibilitar isso, deve-se tornar ambos os parâmetros de codificação equivalentes ao máximo. Por exemplo, o quadro de maior qualidade visual – e consequentemente o que mais faz uso de profundidades baixas da árvore – é o chamado quadro Intra. Esse quadro específico aparece deterministicamente em um período de tempo pré-fixado no HEVC, sendo sempre o primeiro no que ele denomina *Group of Pictures* (GoP). Similarmente ocorre no AV1, com a diferença de que esse GoP não possui uma periodicidade fixa, sendo necessário forçar que haja essa periodicidade. Isso é possível determinando a distância mínima e máxima do *KeyFrame*, nome do quadro Intra do AV1. Logo, optou-se pela configuração mais simples do HEVC, o “*low_delay_main*”, cujo GoP é de quatro quadros.

Outro fator que determina o nível máximo de profundidade dos codificadores, assim como na qualidade final da codificação, é o *Quantization Parameter* (QP). Quanto maior o valor do QP, maior será a perda de qualidade em relação ao vídeo original, possibilitando uma maior compressão de dados. Tanto o HEVC como o AV1 possuem aplicações diferentes para um mesmo valor de QP. Desta forma, selecionou-se um valor tradicional utilizado pelos trabalhos com o HEVC no que se refere ao QP de alta qualidade, ou seja, o QP 22. No outro lado, optou-se por utilizar um valor de QP equivalente à quinta parte do alcance máximo permitido pelo AV1: QP 12.

Como cada padrão possibilita a utilização de blocos de tamanhos e formatos diferentes em cada nível da árvore quadrática, a Tabela 1 apresenta a definição

dos tamanhos de CU no HEVC e dos tamanhos de blocos no AV1 para os níveis de profundidade (DL, do inglês *Depth Level*) possíveis em ambos os padrões.

Para a realização dos testes, foram escolhidos aleatoriamente oito vídeos de resolução 4K disponibilizados pela XIPH (2018). São eles: *Netflix_Boat*, *Netflix_BoxingPractice*, *Netflix_DinnerScene*, *Netflix_Narrator*, *Netflix_RitualDance*, *Netflix_ToddlerFountain*, *Netflix_TunnelFlag* e *Netflix_WindAndNature*. Foram extraídos os quadros 121 a 181 de cada um deles. Foram utilizados os softwares referências do HEVC (HM, 2018), versão 16.18, e do AV1 (AOM, 2018), versão 1.0.

Tabela 1. Definição de Níveis de Profundidade

Profundidade	CUs possíveis (HEVC)	Blocos possíveis (AV1)
0	-	128x128, 128x64, 64x128
1	64x64	64x64, 64x32, 32x64, 64x16, 16x64
2	32x32	32x32, 32x16, 16x32, 32x8, 8x32
3	16x16	16x16, 16x8, 8x16, 16x4, 4x16
4	8x8	8x8, 8x4, 4x8
5	-	4x4

3. RESULTADOS E DISCUSSÃO

A ideia principal da estratégia de transcodificação que seguirá a análise apresentada neste resumo consiste em extrair a árvore de CUs do HEVC durante o processo de decodificação, modificar a forma como esses dados são representados – de código interno de CU para nível de profundidade apresentada na Tabela 1 – e posteriormente limitar a recodificação AV1. Antes disso, contudo, a análise de correlação apresentada nesta seção é essencial para verificar a viabilidade da estratégia e para justificá-la.

Para realização da análise de correlação, a estrutura de blocos foi exportada durante a decodificação do vídeo no AV1, e modificada para se assemelhar à informação obtida pelo codificador HEVC. Uma comparação simples é realizada de forma a se obter uma correlação de níveis de profundidade. Conforme descrito na seção anterior, o codificador AV1 foi configurado para se assemelhar o máximo possível ao HEVC em termos de GoP. Com isso, foi possível obter a Tabela 2, onde é possível observar que, considerando em uma mesma região do vídeo, cada profundidade observada na decodificação HEVC corresponde na maioria das vezes à mesma profundidade decidida na codificação AV1. Por exemplo, em 61,15% das vezes em que a profundidade 2 (DL2) foi utilizada na codificação HEVC, o codificador AV1 escolheu também utilizar DL2 (isto é, os tamanhos de blocos 32x32, 32x16, 16x32, 32x8 ou 8x32).

Conforme previamente abordado, o objetivo futuro desta pesquisa consiste no desenvolvimento de uma limitação inferior de profundidades durante a recodificação no AV1, ou seja, proibir que este codificador utilize tamanhos de blocos menores que aqueles observados no HEVC. Assim sendo, ao somarem-se as colunas até o nível de profundidade observado na decodificação HEVC, pode-se observar uma taxa de acerto de 72,89% para a profundidade DL1, de 84,20% para DL2, de 90,90% para DL3 e de 93,95% para DL4. Os experimentos permitem concluir que a estratégia possui, portanto, boas chances de sucesso, dada a sua baixa taxa de erros.

Tabela 2. Correlação entre Níveis de Profundidade entre HEVC e AV1

HEVC	AV1					
	DL0	DL1	DL2	DL3	DL4	DL5
DL1	25,28	47,61	22,71	4,21	0,19	0,01
DL2	5,76	17,29	61,15	14,49	1,25	0,07
DL3	2,16	5,09	20,76	62,89	8,34	0,76
DL4	0,91	1,94	6,63	39,42	45,06	6,05

4. CONCLUSÕES

Como apresentado neste artigo, a estratégia de possibilitar uma transcodificação tandem mais inteligente que a recodificação completa, utilizando informações de particionamento HEVC para limitar o particionamento de blocos do AV1, possui correlação positiva. Os resultados indicam que há forte correlação e, portanto, fortes indícios que a estratégia pode trazer redução de tempo de transcodificação com baixo impacto na eficiência de codificação. Taxas de acerto de 72,89% para a profundidade DL1, de 84,20% para DL2, de 90,90% para DL3 e de 93,95% para DL4 foram atingidas.

Com isso, pretende-se realizar a implementação de uma transcodificação que utilize as informações de profundidades de CU para limitar diretamente a profundidade máxima dos blocos durante a codificação AV1, reduzindo o tempo de codificação com pequenas perdas na eficiência de compressão.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- AKYAZI, P.; EBRAHIMI, T.. “Comparison of compression efficiency between HEVC/H. 265, VP9 and AV1 based on subjective quality assessments”. In: **10th International Conference on Quality of Multimedia Experience (QoMEX)**. Sardinia, 2018.
- AOM. “**AV1 codec library**”. 2018. Disponível em <https://aomedia.googlesource.com/aom/>
- CHEN, Y.; et. al.. “An overview of core coding tools in the AV1 video codec”, in **2018 Picture Coding Symposium (PCS)**, San Francisco, 2018.
- GROIS, D.; NGUYEN, T.; MARPE, D.. “Performance comparison of AV1, JEM, VP9, and HEVC encoders”. In: **Applications of Digital Image Processing XL**. International Society for Optics and Photonics, 2018. p. 103960L.
- HM. “**HEVC Software**”. 2018. Disponível em https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/
- SULLIVAN, G. J; et al. “Overview of the High Efficiency Video Coding (HEVC) Standard”, in **IEEE Transactions on Circuits and Systems for Video Technology**, vol. 22, no. 12, pp. 1649-1668, 2012.
- WIEGAND, T.; et al. “Overview of the H.264/AVC video coding standard”, in **IEEE Transactions on Circuits and Systems for Video Technology**, vol. 13, no. 7, pp. 560-576, 2003.
- XIPH. “**Xiph.org Video Test Media**”. 2018. Disponível em <https://media.xiph.org/video/derf/>