

ANÁLISE DE EFICIÊNCIA DE CODIFICAÇÃO E CUSTO COMPUTACIONAL DOS MODOS DE SELEÇÃO DE TRANSFORMADA MÚLTIPLA NO CODIFICADOR DE VÍDEO VVC

CAROLINE SOUZA CAMARGO; BIANCA SILVEIRA; GUILHERME CORREA

Universidade Federal de Pelotas (UFPEl) – {caroline.sc, bscsilveira, gcorrea}@inf.ufpel.edu.br

1. INTRODUÇÃO

De acordo com um relatório da Ericsson, 60% do tráfego de dados em smartphones em 2022 foi direcionado para vídeos, e essa proporção deve subir para 74% até 2024 (JEJDLING, 2018). A crescente demanda por conteúdo em alta definição apresenta desafios significativos para as empresas de *streaming*, que se veem na necessidade de otimizar a taxa de bits dos vídeos. Nesse contexto, codecs como o *Versatile Video Coding* (H.266/VVC) se tornam essenciais na transmissão eficiente de conteúdo digital. O VVC oferece uma compressão 50% mais eficiente em relação ao seu predecessor, H.265/HEVC (FRAUNHOFER, 2020). No entanto, é importante mencionar que, para atingir essas altas taxas de compressão, várias novas ferramentas e funcionalidades foram implementadas, resultando em maior poder computacional e consumo de energia.

Entre as novas técnicas introduzidas pelo H.266/VVC destaca-se a *Multiple Transform Selection* (MTS), que tem como objetivo otimizar o desempenho no processo de codificação, especialmente na etapa de transformação. No VTM, *software* de referência do VVC, existem duas variantes da MTS: MTS explícito e MTS implícito. No modo explícito, a seleção dos tipos de transformação é realizada por meio de um processo de sinalização, onde o codificador informa explicitamente qual transformação será aplicada aos dados. Isso proporciona um controle mais direto sobre o tipo de transformação utilizada, permitindo ao codificador adaptar-se a diferentes características do conteúdo.

Por outro lado, no modo implícito, a necessidade de sinalização é eliminada. Nesse caso, as transformações são escolhidas automaticamente com base em informações pré-determinadas que são compartilhadas entre o codificador e o decodificador. Essa abordagem simplifica o processo de codificação, já que as transformações são selecionadas sem a necessidade de comunicação adicional. Vale ressaltar que o MTS implícito é restrito ao contexto de predição intra, onde as transformações adicionais são aplicadas somente a blocos codificados que utilizam este tipo de predição.

A Tabela 1 apresenta as combinações potenciais entre o MTS explícito e implícito. A escolha do modo MTS a ser utilizado durante a codificação pode ser configurada no conjunto de parâmetros da sequência de vídeo presente no VVC. Por sua vez, a Tabela 2 demonstra que, quando o MTS está desabilitado, a transformada aplicada é a DCT-II, tanto na horizontal quanto na vertical. Já quando o MTS está habilitado, novas flags (*Hor flag* e *Ver flag*) indicam o tipo de transformada a ser utilizada e diferentes combinações podem ser escolhidas. Neste trabalho, exploraremos o impacto na eficiência de compressão de diferentes modos e combinações de MTS implícito e explícito utilizados no codificador. Além disso, buscaremos entender qual desses modos apresenta maior custo computacional, proporcionando, assim, oportunidades de otimização.

Modo MTS	Intra	Inter
1	Explícito	-
2	Implícito	Explícito
3	Explícito	Explícito
4	Implícito	-

Tabela 1. Aplicação de MTS explícito e implícito em modos MTS

MTS	Hor flag	Ver flag	1º Transformada	2º Transformada
Desabilitado	-	-	DCT-II	DCT-II
Habilitado	0	0	DST-VII	DST-VII
Habilitado	0	1	DCT-VIII	DST-VII
Habilitado	1	0	DST-VII	DCT-VIII
Habilitado	1	1	DCT-VIII	DCT-VIII

Tabela 2. Flags MTS e combinações de transformações

2. METODOLOGIA

Para analisarmos o impacto dos diferentes modos MTS na eficiência de compressão no codec VVC, foram testadas quatro configurações de codificação: MTS = 1 (Intra explícito), MTS = 2 (Intra implícito e Inter explícito), MTS = 3 (Intra explícito e Inter explícito) e MTS = 4 (Intra implícito). Para realizar os experimentos e coletar dados, utilizamos o software de referência VVC Test Model (VTM), versão 23.0, configurado no modo *Random Access*. Os testes foram aplicados a todos os vídeos do conjunto de referência sugerido por (BOSSSEN, 2013), abrangendo diversos cenários de codificação, o que permitiu avaliar o desempenho do codificador em diferentes condições de textura e movimento. Cada experimento foi conduzido nos primeiros 60 quadros de cada vídeo, utilizando quatro valores de parâmetros de quantização (QP): 22, 27, 32 e 37. Essa abordagem permitiu avaliar o desempenho do codificador sob uma ampla gama de condições, variando de alta a baixa qualidade de imagem.

3. RESULTADOS E DISCUSSÃO

Os resultados obtidos nos experimentos estão apresentados na Tabela 3, que avalia a eficiência e o tempo de codificação em relação à codificação de referência. A codificação de referência utilizada para comparação foi o modo MTS igual a 0, ou seja, com o MTS desabilitado. As colunas BD-Rate e tempo foram organizadas para refletir os diferentes modos MTS. Para facilitar a visualização dos resultados, os vídeos foram agrupados pelas respectivas classes, calculando-se a média de BD-Rate e tempo para cada grupo. A métrica BD-Rate é particularmente relevante nesta análise, pois oferece uma medida quantitativa da eficiência da codificação (BJØNTEGAARD, 2001). Valores negativos indicam uma melhoria na eficiência de codificação e valores positivos indicam uma perda de eficiência de codificação em relação ao caso âncora de comparação. Ao

examinar as variações de BD-Rate entre os diferentes modos MTS, podemos identificar quais configurações proporcionam melhorias significativas em relação à codificação de referência. A análise do tempo de codificação é igualmente importante, pois revela o custo computacional associado à implementação de cada modo MTS.

	MTS = 1		MTS = 2		MTS = 3		MTS = 4	
Classe	BD-Rate (%)	Δ Tempo (%)	BD-Rate (%)	Δ Tempo (%)	BD-Rate (%)	Δ Tempo (%)	BD-Rate (%)	Δ Tempo (%)
A1	-0,58	3,60	-0,37	7,84	-0,59	12,57	-0,31	1,76
A2	-0,89	7,24	-0,95	8,09	-1,15	15,09	-0,63	0,82
B	-0,96	8,06	-0,76	6,20	-1,02	13,73	-0,67	0,46
C	-0,53	11,14	-0,46	6,87	-0,65	17,37	-0,36	0,50
D	-0,52	10,03	-0,48	6,06	-0,48	15,16	-0,39	0,46
E	-1,21	5,50	-0,89	1,39	-1,23	7,83	-0,95	-0,23
F	-0,27	9,28	0,13	3,86	-0,30	12,11	0,13	0,73
Média	-0,71	7,84	-0,54	5,76	-0,77	13,41	-0,45	0,64

Tabela 3. Eficiência e Tempo de Codificação para os Diferentes Modos MTS

Observa-se que, como esperado, os modos MTS trazem melhorias na eficiência da codificação, uma vez que a maioria dos valores de BD-Rate na Tabela 3 é negativa. Essa tendência é esperada, pois a utilização de múltiplas transformações durante a codificação permite uma melhor adaptação às características específicas dos blocos. Assim, o MTS pode aprimorar a precisão da representação de dados, resultando em vídeos de melhor qualidade com uma taxa de bits reduzida, o que indica uma melhoria na compressão final. Notavelmente, o maior benefício em termos de BD-Rate foi observado no modo MTS=3, que apresentou uma média de -0,77% de BD-Rate. Isso sugere que a utilização de transformação explícita em predição intra e inter é particularmente eficaz na compressão de vídeo, melhorando a eficiência da qualidade final. Contudo, embora o MTS=3 tenha proporcionado ganhos na BD-Rate, isso ocorreu com um aumento médio de 13,41% no tempo de codificação, destacando a compensação entre eficiência de codificação e tempo de processamento ao utilizar MTS explícito.

Além disso, considerando que esse modo apresenta o maior consumo computacional e a melhor eficiência de codificação, ele também representa uma oportunidade valiosa para otimizações em trabalhos futuros. O esforço computacional exigido permite explorar melhorias significativas no tempo de codificação. Em contraste, o uso de MTS implícito requer menos esforço do codificador. No caso do MTS=4, o tempo de codificação aumentou apenas 0,64% em relação à codificação original, enquanto houve uma melhoria de -0,45% na eficiência da codificação. Por outro lado, na configuração com MTS=2, onde o MTS implícito foi aplicado na predição intra e MTS explícito na predição inter, o tempo de codificação aumentou consideravelmente para 5,76%. Esse resultado demonstra que o MTS explícito é o modo que demanda maior custo computacional. Em relação à melhoria na codificação, observa-se que o aumento na eficiência não correspondeu proporcionalmente ao tempo de codificação; a

taxa de BD foi de -0,45% com MTS=4 e passou para -0,54% com MTS=2, representando um aumento de apenas -0,09%.

Em resumo, isso indica que o esforço computacional ao ativar o MTS explícito na predição inter é substancialmente alto em relação ao modesto ganho em BD-Rate. Conclui-se, portanto, que os modos MTS que proporcionam um bom equilíbrio entre eficiência de compressão e tempo de codificação são os modos MTS=1 e MTS=2. Essas configurações intermediárias são particularmente significativas, ressaltando a importância de considerar as compensações entre eficiência e recursos computacionais ao optar pelo modo MTS. Essas opções oferecem um meio-termo que equilibra diferentes requisitos de aplicação, proporcionando uma abordagem adaptável e eficaz para a codificação de vídeo em uma variedade de cenários.

4. CONCLUSÕES

Com o aumento da demanda por vídeos em alta definição, as empresas de *streaming* enfrentam o desafio de fornecer conteúdo de qualidade sem comprometer a largura de banda. Nesse cenário, codecs como o H.266/VVC se destacam por sua capacidade de alcançar altas taxas de compressão sem perda significativa de qualidade. No entanto, a implementação desses codecs envolve desafios relacionados à complexidade computacional e ao consumo de energia. A análise dos modos MTS no H.266/VVC mostra que, embora alguns ofereçam melhorias expressivas na eficiência de compressão, isso vem acompanhado de um aumento no tempo de codificação. O MTS explícito, por exemplo, embora aumente a eficiência, tende a exigir mais tempo de processamento em comparação ao MTS implícito. Em contrapartida, modos como MTS=1 e MTS=2 proporcionam um equilíbrio mais favorável entre compressão e tempo de codificação, tornando-os opções viáveis para um meio-termo eficiente.

5. REFERÊNCIAS BIBLIOGRÁFICAS

JEJDLING, F. **Ericsson mobility report**. Nov. 2018. Disponível em: <https://www.ericsson.com/en/reports-and-papers/mobility-report/reports>. Acesso em 24 de setembro.

FRAUNHOFER INSTITUTE FOR TELECOMMUNICATIONS. **H.266/VVC brings video transmission to new speed**. Jul. 2020. Disponível em: <https://newsletter.fraunhofer.de/-viewonline2/17386/465/11/14SHcBTt/V44RELLZBp/1>. Acesso em 24 de setembro.

BOSSEN, F. **Common test conditions and software reference configurations**. April 2013. Disponível em: <http://phenix.it-sudparis.eu/jct/doc/end/user/current/document.php?id=7281>. Acesso em: 24 de setembro.

BJØNTEGAARD, G. **Calculation of average PSNR differences between RD-curves**. 2001. Disponível em: <http://wftp3.itu.int/av-arch/video-site/0104/Aus/VCEG-M33.doc>. Acesso em: 24 de setembro.