

CONTROLADOR DE COMPLEXIDADE PARA O CODIFICADOR DE VÍDEO AV1

GUSTAVO HÜBNER REHBEIN¹; ISIS DUARTE BENDER^{1,2}; GUILHERME
CORRÊA¹; LUCIANO AGOSTINI¹; MARCELO SCHIAVON PORTO¹

¹Universidade Federal de Pelotas – Video Technology Research Group (ViTech)

²Instituto Federal Sul-rio-grandense Campus Pelotas
{ghrehbein, idbender, gcorrea, agostini, porto} @inf.ufpel.edu.br

1. INTRODUÇÃO

Nos últimos anos, a existência de dispositivos modernos e conexões de internet de alta velocidade contribuíram, significativamente, com o crescimento das aplicações e serviços de vídeos digitais. Atualmente, o Youtube possui, cerca de 2 bilhões de usuários ativos mensais (YOUTUBE, 2022).

Vídeos digitais apresentam um grande volume de dados. Para que a transmissão e armazenamento desse tipo de mídia seja viável, é necessário o uso de compressores de vídeo. Existem vários padrões e formatos de codificação disponíveis, os quais buscam entregar uma maior eficiência de compressão em relação a codificadores anteriores. Dentre eles, o formato de codificação AOMedia Video 1 (AV1) é um dos mais recentes. Lançado em 2018 pela *Alliance for Open Media* (AOMedia) (AOMEDIA, 2018). O AV1 foi desenvolvido com base no seu antecessor, o codificador VP9, trazendo novas ferramentas e aprimoramentos para alcançar uma maior eficiência de codificação. Porém, esses aprimoramentos aumentaram o custo computacional do AV1 consideravelmente, quando comparado com o VP9 (MANSRI et al., 2020).

Na literatura, é possível encontrar diversas soluções publicadas propondo estratégias para diminuir a complexidade de codificação do AV1. Porém, ser capaz de controlar sua complexidade dinamicamente também pode ser interessante, pois permite que a complexidade seja adequada de acordo com as limitações e requisitos impostos pelo sistema em que ele está sendo executado. Isso é benéfico especialmente para dispositivos movidos a bateria, como *smartphones* e *tablets*.

Neste trabalho é apresentado um controlador de complexidade desenvolvido para o codificador AV1 baseado na otimização multi-objetivo. O controlador apresentado pode reduzir a complexidade de codificação com alvos variando de 10% a 40% de redução de tempo de codificação, com uma precisão variando de 0,04 a 4,1 pontos percentuais e um aumento de BD-Rate (BJONTEGAARD, 2001) variando de 0,22% a 5,24% para os vídeos e alvos de tempo avaliados.

2. METODOLOGIA

O algoritmo de controle de complexidade apresentado neste trabalho explora a seleção dos parâmetros de codificação do AV1 a partir de uma otimização multi-objetivo. A fim de avaliar o impacto de cada configuração sobre o tempo de codificação e eficiência de codificação, foram realizados experimentos testando o impacto de cada parâmetro de codificação do AV1. Uma análise prévia das etapas de codificação do AV1 demonstrou que as etapas inter-quadros e transformadas possuem a maior complexidade computacional, correspondendo a cerca de 95% do tempo total de codificação. Assim, os parâmetros de codificação referentes a estas duas etapas foram selecionados para serem avaliados.

Nos experimentos foram utilizadas 10 seqüências de teste, sendo 5 HD 1080p e 5 UHD 4K. Os primeiros 60 quadros foram codificados usando o software de referência Libaom (GOOGLEGIT, 2022). Foi utilizado a configuração *High Latency CQP* com os CQs 20, 32, 43 e 55 conforme recomendado em DAEDE; NORKIN; BRAILOVSKIY (2022). Ao todo foram avaliados 38 parâmetros de configuração do AV1, os resultados mostraram que os parâmetros conseguem entregar uma redução de tempo de codificação variando entre -0,35% e 79.38% com um aumento de BD-Rate variando entre 0,05 e 69,83.

Para aprimorar os efeitos de redução de tempo e eficiência de codificação, esses parâmetros foram combinados em conjunto de configurações. Foi realizada uma análise de Frente de Pareto (NGATCHOU; ZAREI; EL-SHARKAWI, 2005) para selecionar as configurações com a melhor relação entre redução de tempo e BD-Rate. Assim, foram selecionadas 6 configurações, as quais foram combinadas incrementalmente, gerando 15 conjuntos de configurações.

Uma nova análise de Frente de Pareto, considerando as configurações individuais e os conjuntos de configurações testados, foi aplicada para selecionar 7 pontos de controle, além do ponto que utiliza a configuração padrão do AV1. Essas configurações podem entregar uma redução do tempo de codificação variando de 0%, para a configuração padrão, até 39,96% para a configuração com a maior redução de tempo de codificação.

O controlador de complexidade apresentado neste trabalho está ilustrado na Figura 1 (a). A cada 16 quadros, um tempo alvo é informado ao controlador. Esse tempo é comparado com o tempo medido nos últimos 16 quadros codificados para, então, realizar a seleção de uma configuração que entregue a redução de tempo necessária para alcançar o tempo alvo informado. A seleção da configuração é feita a partir do uso de uma tabela de transição apresentada na Figura 1 (b), onde as colunas indicam a última configuração usada e as linhas indicam a complexidade relativa das outras configurações em relação à configuração atual.

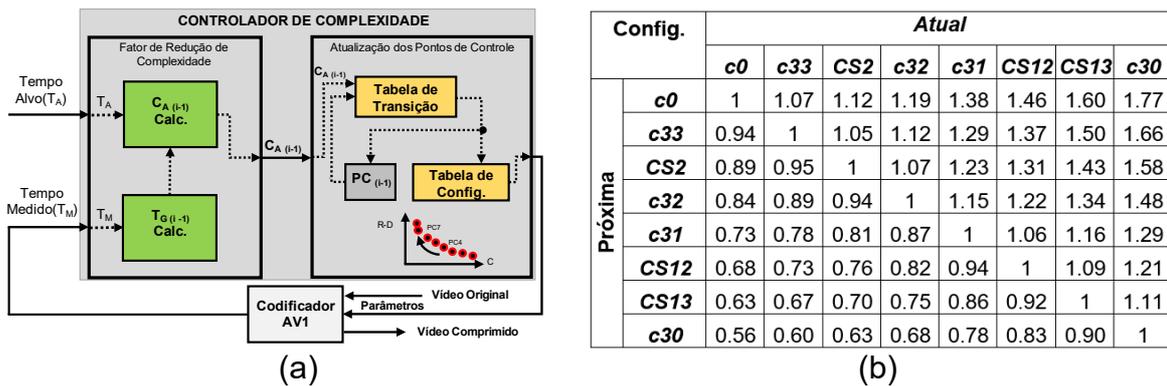


Figura 1 – Funcionamento do Controlador (a) e Tabela de Transição (b).

3. RESULTADOS E DISCUSSÃO

O controlador de complexidade apresentado foi avaliado usando 3 seqüências de teste em HD 1080 e 3 seqüências em UHD 4K, cada uma com 180 quadros. Para esses experimentos foram usados os CQs 20, 32, 43 e 55, com a redução do tempo de codificação fixada em 10%, 20%, 30% e 40%. A versão do Libaom utilizada foi a versão 2.0 – hash 72824a7. O computador utilizado possuía um Intel Xeon E5-4650v3v4 de 2,1 GHz com 512 GB de memória RAM.

Os resultados desses experimentos são apresentados na Tabela 1, onde são apresentados os resultados médios para as seis sequências de testes utilizadas. É possível perceber que os valores obtidos se mantiveram próximos do alvo de redução de tempo especificado para cada experimento, com um desvio padrão variando de 0,62 a 1,71 pontos percentuais. Os resultados para os experimentos de 40% de redução apresentam a maior diferença entre o alvo especificado e o resultado obtido, o que era esperado devido à proximidade com o valor máximo de redução que pode ser obtido com as configurações selecionadas. Os resultados de aumento de BD-Rate obtidos variam entre 0,44% a 5,14%, com os maiores valores nos experimentos de maior redução de tempo.

Tabela 1 – Resultados dos Experimentos.

Sequência de Teste		Red. de Tempo de Codificação			
		10%	20%	30%	40%
Red. Média Obtida	ducks_take_off	10,84%	21,63%	29,13%	37,38%
	park_joy	11,17%	19,68%	27,48%	34,77%
	NF_FoodMarket	11,40%	20,74%	28,70%	37,13%
	NF_Narrator	9,89%	20,90%	28,23%	36,11%
	NF_WindAndNature	9,83%	22,11%	28,63%	36,13%
	NF_Dancers	10,17%	21,00%	30,44%	35,47%
	Média	10,55%	21,01%	28,77%	36,16%
Desvio Padrão (p.p.)		0,62	0,76	0,90	1,71
Aumento de BD-Rate		0,44%	2,68%	2,99%	5,14%

A Figura 2 mostra o resultado do uso do controlador na sequência *Netflix_Dancers* com o CQ 55 e redução de tempo de codificação em 40%. O gráfico mostra os resultados do tempo de codificação por grupo de 16 quadros, com e sem o uso do controlador, o tempo médio de codificação sem o controlador e o tempo alvo definido. Também é apresentada a configuração selecionada pelo controlador para cada grupo. É possível perceber que o controlador conseguiu estabilizar o tempo de codificação próximo ao definido no tempo alvo.

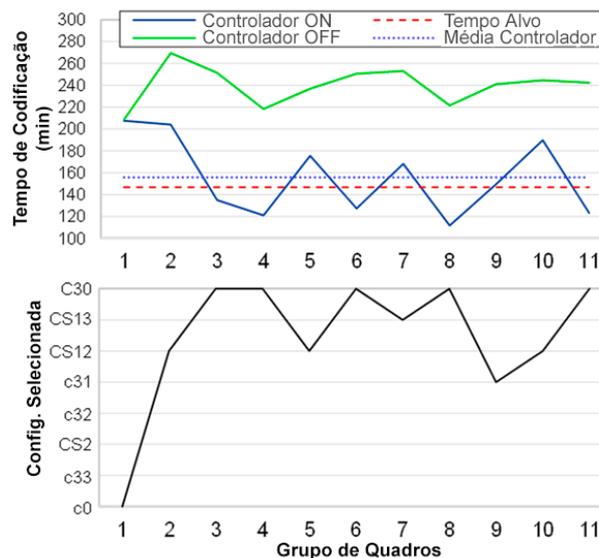


Figura 2 – Comportamento do Controlador.

4. CONCLUSÕES

Esse trabalho apresentou um controlador de complexidade baseado na otimização multi-objetivo para o AV1. Uma ampla análise de sensibilidade dos parâmetros de codificação do Libaom foi realizada para identificar seu impacto na redução do tempo de codificação e na eficiência da codificação. A frente de Pareto foi utilizada para identificar os parâmetros mais relevantes e criar um conjunto de pontos de controle a serem utilizados no controlador.

O controlador foi desenvolvido considerando uma granularidade de atualização de 16 quadros e foi avaliado para sequências de vídeo HD e UHD, com reduções de tempo alvo entre 10% e 40%. Em uma aplicação real, o fator de redução de complexidade será uma entrada externa ao controlador e codificador, que pode vir do usuário externo ou do próprio sistema operacional.

Resultados experimentais mostram que o controlador atinge a meta de redução de tempo com uma precisão variando de 0,10 a 5,20 pontos percentuais e um aumento de BD-Rate de 0,44% a 5,14%. Como trabalho futuro, é esperado melhorar o controlador explorando as características das sequências de vídeo, decisões de codificação e métodos adicionais de controle fino para minimizar ainda mais as perdas de BD-Rate.

5. REFERÊNCIAS BIBLIOGRÁFICAS

AOMEDIA. **AV1 Bitstream and decoding process specification**. 2018. Acessado em abril de 2022. Online. Disponível em: <https://aomediacodec.github.io/av1-spec/av1-spec.pdf>

BJONTEGAARD, G. Calculation of average PSNR differences between RD-curves. **VCEG-M33**, [S.l.], 2001.

DAEDE, T.; NORKIN, A.; BRAILOVSKIY, I. **Video Codec Testing and Quality Measurement draft-ietf-netvc-testing-09**. 2022, Amazon Lab126. ago. 2020. p. 23. Acessado em abril de 2022. Online. Disponível em: <https://tools.ietf.org/html/draft-ietf-netvc-testing-09>

GOOGLEGIT. **AV1 Codec Library**. 2022. Acessado em abril de 2022. Online. Disponível em: <https://aomedia.googlesource.com/aom/>

MANSRI, I. et al. Comparative Evaluation of VVC, HEVC, H.264, AV1, and VP9 Encoders for Low-Delay Video Applications. In: **FOURTH INTERNATIONAL CONFERENCE ON MULTIMEDIA COMPUTING, NETWORKING AND APPLICATIONS (MCNA)**, 2020., 2020. **Anais**. . . [S.l.: s.n.], 2020. p.38–43.

NGATCHOU, P.; ZAREI, A.; EL-SHARKAWI, A. Pareto multi objective optimization. In: **OF THE 13TH INTERNATIONAL CONFERENCE ON, INTELLIGENT SYSTEMS APPLICATION TO POWER SYSTEMS**, 2005. **Proceedings**. . . [S.l.: s.n.], 2005. p.84–91.

YOUTUBE. **YouTube for Press**. 2022. Acessado em abril de 2022. Online. Disponível em: <https://blog.youtube/press/>