

ANÁLISE DA UTILIZAÇÃO DAS FERRAMENTAS PARA CODIFICAÇÃO DE CONTEÚDO DE TELA NO CODIFICADOR DE VÍDEO HEVC-SCC

EDUARDO DIAS DA SILVA; MARCELO PORTO; GUILHERME CORRÊA

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

1. INTRODUÇÃO

Atualmente, um tipo de vídeo capturado, não por câmeras, mas diretamente da tela do computador, vem ganhando popularidade e se tornando cada vez mais comum. Estes são os chamados vídeos de conteúdo de tela (*Screen Content – SC*). *Displays* sem fio, *webinars*, videoconferência com compartilhamento de tela, streaming de jogos na web, entre outras aplicações, têm como principal fonte de conteúdo este tipo de vídeo.

Conforme (XU; JOSHI; COHEN, 2016), os vídeos utilizados nestes tipos de aplicação podem conter quantidades significativas de SC, que normalmente são compostas por regiões significativas que permanecem estáticas. Quando há movimento, este é bastante homogêneo e contínuo, pois corresponde a um conjunto de caracteres (textos) e imagens sintéticas. Em poucos casos, imagens realmente capturadas por câmeras são utilizadas. Geralmente, estas ocupam uma pequena parcela da área total do vídeo.

Em 2015, o *Twitch* (TWITCH, 2017), plataforma especializada em transmissão de jogos eletrônicos na web, alcançou a marca de 241 bilhões de minutos transmitidos. Em 2016, foram 292 bilhões de minutos (TWITCH, 2017). Este tipo de transmissão gera uma grande quantidade de tráfego na rede, de forma que algoritmos e padrões de compressão de vídeo especializados para este tipo de conteúdo são extremamente necessários para reduzir o tráfego na rede.

Em janeiro de 2016, foi lançada a extensão *Screen Content Coding HEVC* (SCC-HEVC), uma extensão do padrão *High Efficiency Video Coding* (HEVC) dedicada exclusivamente à codificação de SC (XU; JOSHI; COHEN, 2016). Diferentes das imagens capturadas por câmeras, vídeos de SC geralmente não contêm ruídos e contêm grandes áreas homogêneas, padrões de blocos repetidos, quantidades reduzidas de cores, além de regiões idênticas durante diversos quadros consecutivos. Explorando essas características, o SCC-HEVC busca obter uma maior eficiência de compressão de vídeos SC do que o HEVC puro, que foi construído com foco em vídeos de propósito genérico (majoritariamente capturados por câmeras).

O SCC-HEVC introduz quatro novas ferramentas ao padrão HEVC, sendo estas chamadas *Intra-Block Copy* (IBC), *Palette Mode* (PM), *Adaptive Color Transform* (ACT) e *Adaptive Motion Vector Resolution* (AMVR).

O IBC é responsável por explorar redundâncias entre blocos muito semelhantes dentro de um mesmo quadro, associando vetores de movimento a estes. Este tipo de predição é similar à predição inter-quadros no HEVC, entretanto ocorre dentro do mesmo quadro (e não entre quadros temporalmente vizinhos). Por isso, apesar do nome *Intra-Block Copy*, este modo é tratado pelo codificador de referência como uma predição do tipo inter. O PM é responsável por explorar a limitada quantidade de cores que compõe cada bloco, associando um índice a cada cor de pixel e codificando estes índices diretamente. A ACT é responsável por

separar as dependências entre cada componente dos canais R, G e B aplicando uma transformada do espaço de cores RGB para o espaço YCoCg, pois em geral a codificação é mais eficiente quando a componente de luminância (Y) é separada das demais, o que não ocorre no espaço RGB, onde cada componente contém uma parte da luminância. Por fim, o AMVR aproveita-se do fato de que muitos blocos utilizam vetores de movimento com valores inteiros para não sinalizar a parte fracionária.

Por possuir todas as ferramentas do HEVC e ainda as suas próprias, o SCC-HEVC possui complexidade superior à do HEVC. Diante do exposto, o objetivo deste trabalho consiste em coletar dados sobre a utilização do conjunto de ferramentas do SCC-HEVC a fim de, futuramente, realizar modificações no codificador com o objetivo de reduzir a sua complexidade em termos de tempo de processamento, sempre focando em manter a eficiência de codificação aceitável.

2. METODOLOGIA

Para a realização deste trabalho, foi utilizado o *software* de referência do HEVC SCC, o *HEVC-SCC Test Model* (HM-SCC) que é formado pelo *HEVC Test Model* (HM-16.15) em conjunto com a extensão *Screen Content Coding Test Model* (SCM-8.4). Inicialmente, foi necessário identificar no código onde são feitas as escolhas dos modos de particionamento e predição utilizados na codificação de vídeo. Logo após, foram inseridas algumas linhas de código que possibilitaram apenas a captura de dados, de modo que essas alterações não causaram impacto nas decisões do codificador. A seguir, um conjunto de oito vídeos no formato RGB foram codificados pelo software HM-SCC (1920x1080: *sc_console*, *sc_desktop*, *sc_flyingGraphics*, *sc_MissinControlClip3*, *sc_socialnetworkMap*; 1280x720: *sc_robot*, *sc_SlideShow* e *sc_web_browsing*). Apenas os 100 primeiros quadros de cada vídeo foram utilizados. Cada vídeo deste conjunto foi codificado em quatro configurações diferentes no modo *Random Access*, com as ferramentas IBC, PM, ACT e AMVR habilitadas, variando apenas o parâmetro de quantização (QP) em cada uma das configurações, conforme indicado por (YU, 2015): 22, 27, 32 e 37.

Os dados capturados durante a codificação foram os seguintes: a quantidade de *Coding Tree Units* (CTUs) codificadas, a quantidade de *Coding Units* (CUs) utilizadas em cada tamanho possível (64x64, 32x32, 16x16 e 8x8), os modos de predição das CUs (inter ou intra), os modos de particionamento das CUs em *Prediction Units* (PUs) (2Nx2N, 2NxN, Nx2N, NxN, 2NxN, 2NxN, nLx2N e nRx2N), o modo de predição escolhidos para as PUs (IBC, IBC Skip, IBC Merge, Inter, Skip, Merge, Intra Planar, Intra DC, Intra Angular e Intra Palette) e a quantidade de blocos de tamanho 4x4 nos quais foi aplicada a ACT. Além destas estatísticas, foram capturados dados referentes ao tempo de processamento de cada uma das novas ferramentas.

Finalmente, os dados coletados foram compilados, classificados e analisados, levando à discussão apresentada na seção a seguir. Ainda no escopo da metodologia, o próximo passo desta pesquisa consiste em realizar decisões baseadas na análise apresentada neste resumo para futuras modificações no codificador de referência (HM-SCC), com o objetivo de redução da sua complexidade.

3. RESULTADOS E DISCUSSÃO

Como a futura solução de redução de complexidade (objetivo-final desta pesquisa) deve ser genérica, ou seja, funcionar independente do vídeo de entrada do codificador, os gráficos apresentados nesta seção foram criados a partir da média das estatísticas obtidas durante a codificação dos 8 vídeos de referência.

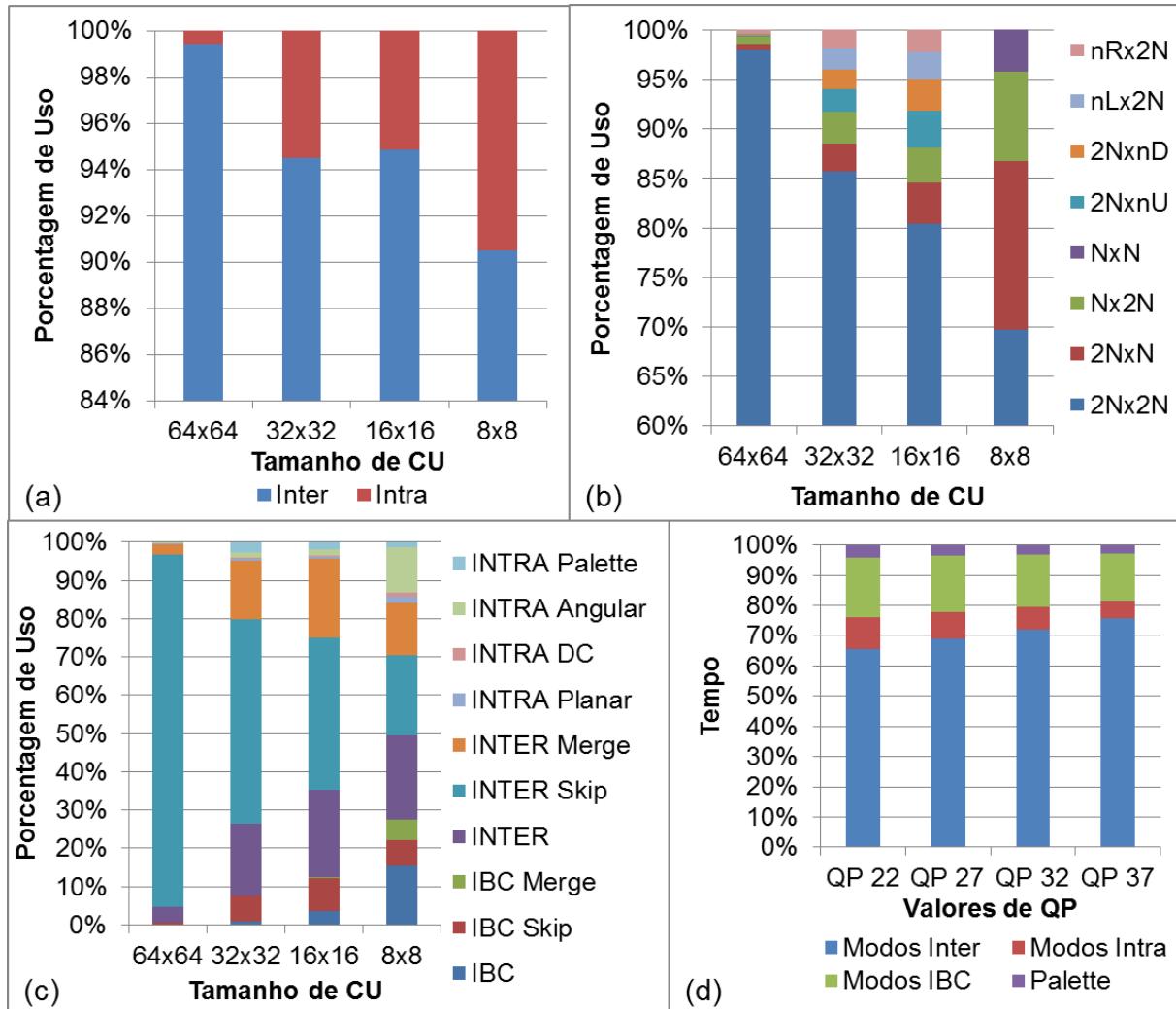


Figura 1: (a) Taxa de ocorrência dos modos inter-quadros e intra-quadro em cada tamanho de CU; (b) Taxa de ocorrência dos diferentes modos de particionamento das CUs em PUs; (c) Taxa de ocorrência dos modos de predição em cada tamanho de CU; (d) Tempo consumido por cada conjunto de modos de predição em cada valor de QP.

Conforme pode ser observado na Figura 1(a), o codificador HM-SCC acabou optando pela predição inter-quadros em 99% das CUs de tamanho 64x64. A utilização do modo *Skip* da predição inter-quadros (chamado *Inter Skip* na Figura 1(c)) tem grande utilização. Isso se dá porque o modo *Skip* é um dos primeiros modos avaliados durante a codificação e, caso ele tenha o melhor custo taxadistorção entre os três primeiros modos avaliados, os modos restantes são automaticamente descartados e o algoritmo termina a busca pelo melhor modo de predição para aquele bloco. O restante das CUs inter-quadros (as que não utilizam o *Inter Skip*) são uma parcela pouco significativa em CUs 64x64. Em CUs menores, a

utilização dos demais modos cresce, mas a concentração ainda se mantém no grupo de modos inter-quadros. Alguns destes modos fazem parte do grupo que é avaliado inicialmente, porém não possuem a restrição de descartar os outros modos e saltar para a próxima etapa do algoritmo caso sejam escolhidos, além dos submodos do IBC (IBC *Merge* e IBC *Skip*).

É importante perceber que, apesar de os sub-modos IBC apresentarem uma parcela pequena de ocorrência, o seu tempo de processamento não é irrelevante. A Figura 1(d) mostra que, em média, cerca de 18% do tempo de codificação acontece em modos IBC, apesar de o somatório destes modos ser escolhido em cerca de 11% das vezes (Figura 1(c)). Portanto, uma das possíveis abordagens para redução de complexidade seria focar na aceleração da decisão de modo IBC, procurando heurísticas que permitam realizar terminações prematuras deste modo em determinadas situações. A Figura 1(d) mostra que os modos *Inter* são os principais responsáveis pela grande complexidade da codificação. Entretanto, diversos outros trabalhos já propõem soluções para reduzir a complexidade da predição inter-quadros. O foco desta pesquisa deve se concentrar, portanto, nos novos modos específicos para conteúdo de tela.

Além disso, uma segunda possível solução para a redução de complexidade seria avaliar nas CUs 64x64 apenas os três modos iniciais (*Inter Skip*, *Inter Merge* e *Inter*), já que os demais praticamente não são utilizados (ver Figura 1(c)). Nos demais tamanhos de CU, uma abordagem interessante seria avaliar primeiro os modos *Inter* e, se necessário, avaliar os modos IBC, evitando os modos *Intra*, raramente utilizados.

4. CONCLUSÕES

Este trabalho apresentou uma análise de utilização das ferramentas, modos de particionamento e tempo utilizado pelos modos de predição do padrão HEVC SCC na codificação de conteúdo de tela. Os resultados mostraram que, após os modos de predição inter-quadros, os modos de predição IBC são os mais complexos, apesar de serem escolhidos poucas vezes pelo codificador. Percebe-se que, com pequenas modificações no codificador HM-SCC, pode ser possível atingir redução de complexidade. Porém, a estratégia precisa ser consistente e inteligente para evitar perdas grandes de eficiência de codificação.

5. REFERÊNCIAS BIBLIOGRÁFICAS

XU, Jizheng; JOSHI, Rajan; COHEN, Robert A. Overview of the emerging HEVC screen content coding extension. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 26, n. 1, p. 50-62, 2016.

TWITCH, **Twitch The 2015 Retrospective**. Twitch. Acessado em: outubro de 2017. Online. Disponível em: <https://www.twitch.tv/year/2015>.

TWITCH, **Presenting the Twitch 2016 Year in Review**. Twitch Blog. Acessado em: outubro de 2017. Online. Disponível em: <https://blog.twitch.tv/presenting-the-twitch-2016-year-in-review-b2e0cdc72f18>.

Yu et al.. **Common Test Conditions for Screen Content Coding**. JCT-VC, Warsaw, 2 dez. 2015. Acessado em 1 out 2017. Online. Disponível em: http://phenix.itsudparis.eu/jct/doc_end_user/ current_document.php?id=10221