MEMORY PROFILING OF H.266 VERSATILE VIDEO CODING STANDARD

ARTHUR CERVEIRA[1]; LUCIANO AGOSTINI[2], BRUNO ZATT[3]; FELIPE SAMPAIO[4]

[1] Universidade Federal de Pelotas (UFPel) – aacerveira@inf.ufpel.edu.br
[2] Universidade Federal de Pelotas (UFPel) – agostini@inf.ufpel.edu.br
[3] Universidade Federal de Pelotas (UFPel) – zatt@inf.ufpel.edu.br
[4] Instituto Federal do Rio Grande do Sul (IFRS) - Campus Farroupilha – felipe.sampaio@farroupilha.ifrs.edu.br

1. INTRODUCTION

In 2018, the Joint Video Experts Team (JVET) started the standardization process for the H.266 Versatile Video Coding (H.266/VVC) (BROSS, 2018), which was defined as a new standard in July 2020. This new video coding standard can compress up to 2x the amount of data in a video, considering the same quality loss compared to its predecessor, the High Efficiency Video Coding (HEVC) (FRAUNHOFER, 2020). To achieve this compression efficiency, H.266/VVC uses novel memory-intensive techniques, requiring up to 13x higher memory bandwidth than HEVC (CERVEIRA, 2020). In this context, it is mandatory specific assessments of the memory requirements of the new H.266/VVC encoders, in order to properly measure such high memory overhead.

Evaluations and optimized techniques targeting memory requirements for video coding tools have been widely proposed for H.264/AVC and HEVC coders in the past years (MATIVI, 2016). These studies highlight the relevance of providing memory-efficient solutions to video codecs. In the context of H.266/VVC, the literature still lacks on addressing these memory aspects, mainly when looking for novel tools proposed by JVET. In (PAKDAMAN, 2020) and (TANOU, 2019), the authors present evaluations and comparisons of the new standard with other video coders, such as HEVC and AV1; however, these works focus on coding efficiency and computational complexity. In (CERVEIRA, 2020), the H.266/VVC memory requirements are compared to HEVC, displaying strong increases in terms of memory accesses, both in overall and inter-prediction perspectives. Even though a memory profiling is performed, the related work does not go deep to analyze ultra-high definition videos (one of H.266/VVC targets). Besides, none of these studies performs an analysis specifically targeting important H.266/VVC novel tools and algorithms present on its encoder. While there are studies targeting such novel coding tools introduced in H.266/VVC (BLÄSER, 2018), they rarely address the issues related to memory usage. Still, *there is an open gap on proper discussions about the memory requirements of these innovations*. Therefore, *the significance of these new tools and algorithms regarding memory usage must be assessed in detail*.

In this sense, the goal of this work is to perform a detailed memory profiling of the H.266/VVC encoder. We intend to provide key remarks regarding the video coder memory accesses, while also proposing research prospects to enable memory-efficient video processing.

As the main contributions of this work, we carry out the H.266/VVC memory analysis through an overall perspective, grouping the memory accesses according to the coding tools at H.266/VVC encoder application. The distribution of the memory requirements is exposed, determining high and low data-intensive tasks. At this point, novel coding tools, mainly related to flexible block partitioning and multiple transforms selections, are discussed.

2. METHODOLOGY

The VTM 8.0 was used as the test model for the H.266/VVC assessment, respecting the common test conditions adopted by the video coding community (CTCs) (JVET, 2020). *Random Access* encoder configuration was selected since it represents the most common prediction structure adopted in real-world multimedia systems. We selected seven video sequences for the experiments: UHD 4K (3840x2160 pixels) from A1 (*Campfire* and *Tango2*) and A2 classes (*DaylightRoad2* and *CatRobot*); and HD1080p (1920x1080 pixels) from B class (*BasketballDrive*, *BQTerrace*, and *Cactus*). The used quantization parameters (QP) were: 37, 32, 27, 22, as defined by the CTCs. The executions were performed in the first 17 frames of each video sequence.

The experiments employed the Intel® VTune™ Profiler for memory profiling (INTEL, 2020). This tool provides the loads (read accesses) and stores (write accesses) counting of each executed function. Then, the functions were mapped to their respective H.266/VVC modules and operations targeted in each analysis.

3. RESULTS AND DISCUSSIONS

This section presents the results of our analyses. Here, we exhibit a memory usage breakdown in each encoding module of the H.266/VVC encoder, considering the impact of variations on the QP value, the video sequences resolution, and its temporal and spatial characteristics.

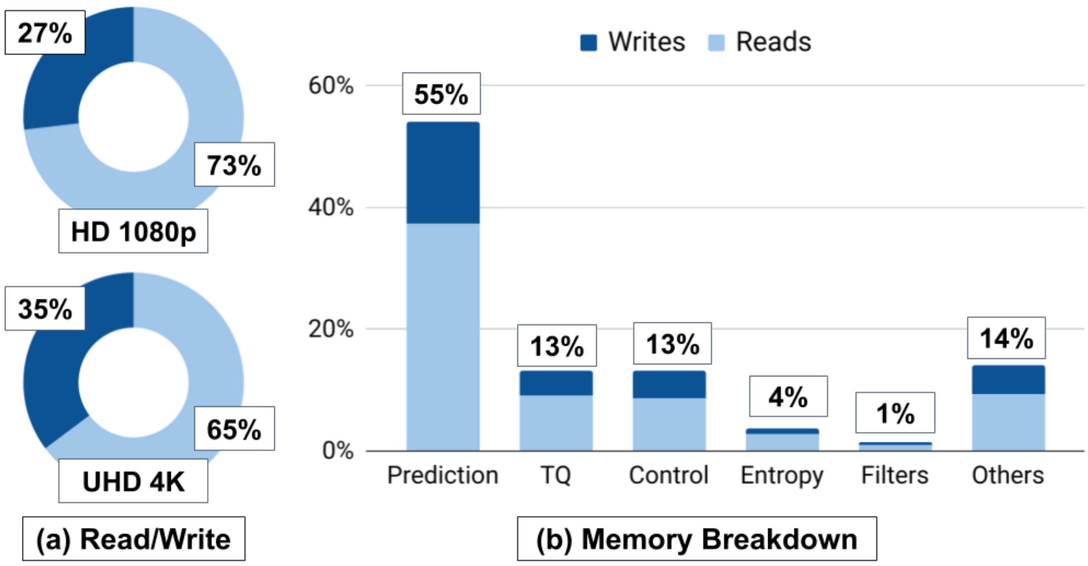

Fig. 1: Encoder memory breakdown: (a) read/write and (b) VVC/H.266 encoding modules.

Figure 1 displays the distribution of memory accesses for the H.266/VVC encoder. For this analysis, we consider the following modules: Prediction, Transform/Quantization (T/Q), Entropy, and Filters. The "Control" step mostly reunites auxiliary functions to manage the encoding flow of the H.266/VVC coding structures: coding-tree units (CTUs), coding units (CUs) and transform units (TUs). Then, we assorted the remaining accesses in the "Others" category. The Prediction module is composed of inter-frame and intra-frame steps. The results of Fig. 1 are achieved by calculating the average read and write accesses for all tested video sequences and QP values, normalized to the total accesses performed during the encoding process.

As one can notice in Fig. 1a, 68% of the memory requirements (on average) are due to read accesses. UHD video sequences exhibit a slightly lower read accesses intensity: 65%, compared to 73% in HD videos. The Prediction module is the most memory-intensive operation in H.266/VVC, representing 55% of all accesses (Fig. 1b). The Prediction kernel employs complex algorithms to find temporal and spatial correlations in a video, such as integer and fractional motion estimation, as well as H.266/VVC-novel tools like the CTU partitioning strategy based on quad-tree with nested binary and ternary trees structures, which explains the significant overhead on its memory requirements. As an initial remark, we observed that *the prediction steps impose the hardest memory challenges during the encoding process, demanding more effective memory optimization strategies to enable efficient H.266/VVC encoding*.

Although the overall distribution of the modules' accesses remains similar, variations on the QP value strongly affect the encoder memory requirements. Fig. 2 presents an evaluation of this impact through absolute and normalized perspectives.

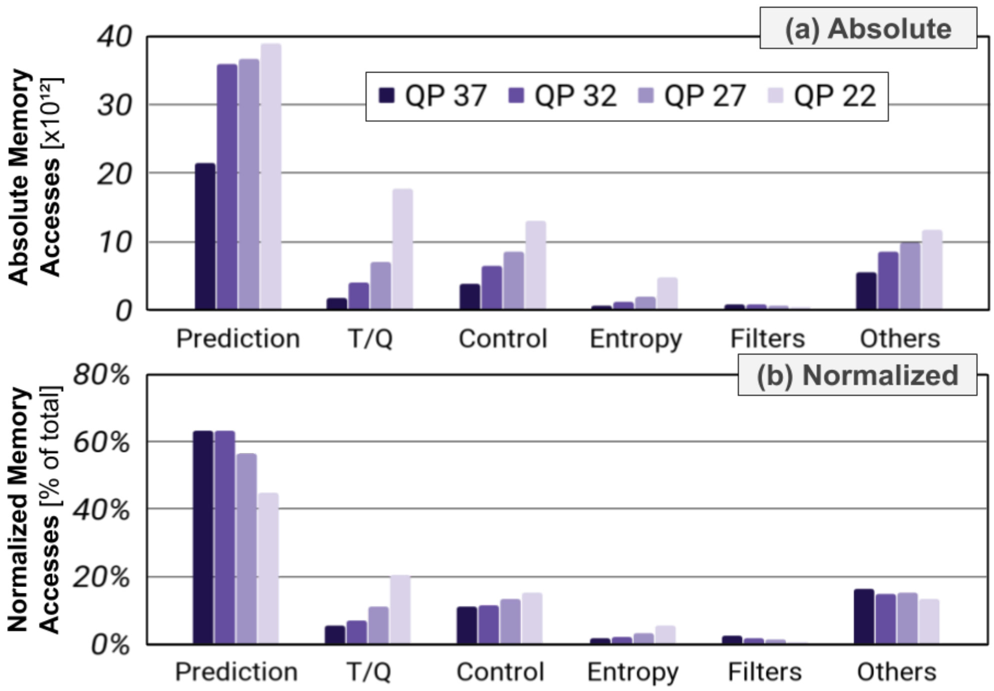

Fig. 2: Encoder memory breakdown: (a) absolute and (b) normalized perspectives.

When the QP gets lower, the quantization step alleviates the cuts over transformed coefficients, requiring higher complexity for the entire prediction mode decision, which affects the memory usage of all encoding modules. The impact of the QP value can be noticed at the absolute results (Fig. 2a). Through this perspective, all encoding modules present significant memory requirements increase as the QP value gets lower. In the normalized results (Fig. 2b), the requirements of each module are obtained through their proportion to the total amount of memory accesses captured during the encoding process. When considering this perspective, the Prediction volume of accesses declines as the QP value gets lower: decreasing from about 63% for QP 37 to 44% for QP 22. Besides this reduction, Prediction remains the most memory-intensive module in all analyzed cases. Thus, another important remark is that *the QP value has a significant influence on the distribution of memory accesses across the encoding modules. Thus, strategies aiming to optimize H.266/VVC memory infrastructure must consider how variations on this parameter affect the memory communication*.

4. CONCLUSIONS

In this work, we presented a profiling of the memory requirements for the H.266/VVC encoder, emphasizing the influence of the prediction steps and new algorithms that were not present in previous coding standards. Through our experiments, we observed how the prediction steps require the highest volume of memory on the encoder, while also acknowledging how the QP value affects this result. This is the first work to perform this sort of analysis, and the results indicate novel perspectives and research prospects on optimizations to reduce the high memory requirements of the H.266/VVC.

5. REFERENCES

BROSS, B. **Working Draft 1 of Versatile Video Coding**. MPEG 122, San Diego, 16 abr. 2018. Acessado em 21 set. 2020. Online. Disponível em: https://mpeg.chiariglione.org/standards/mpeg-i/versatilevideo-coding/n17669-working-draft-1-versatile-video-coding

FRAUNHOFER. **Fraunhofer HHI is proud to present the new state-of-the-art in global video coding: H.266/VVC brings video transmission to new speed**, 6 jul. 2020. Acessado em 21 set. 2020. Online. Disponível em: https://newsletter.fraunhofer.de/-viewonline2/17386/465/11/14SHcBTt/V44RELLZBp/1

CERVEIRA, A.; AGOSTINI, L.; ZATT, B.; SAMPAIO, F. **Memory Assessment of Versatile Video Coding**. IEEE International Conference on Image Processing, p. 1 - 5, 2020.

MATIVI, A.; MONTEIRO, B.; BAMPI, S. **Memory access profiling for HEVC encoders**. IEEE Latin American Symposium on Circuits and Systems, p. 1 - 5, 2016.

PAKDAMAN, F.; ADELIMANESH, M.; GABBOUJ, M.; HASHEMI, R. **Complexity Analysis Of Next-Generation VVC Encoding and Decoding**. IEEE International Conference On Image Processing, p. 1 - 5, 2020.

TANOU, J.; BLESTEL, M. Analysis of Emerging Video Codecs: Coding Tools, Compression Efficiency. **SMPTE Motion Imaging Journal**, v. 128, n. 10, p. 14 - 24, 2019.

BLÄSER, M.; SCHNEIDER, J.; SAUER, J.; WIEN, M. **Geometry-based Partitioning for Predictive Video Coding with Transform Adaptation**. Picture Coding Symposium, p. 1 - 4, 2018.

JVET. **JVET common test conditions and software reference configurations for SDR video**, 2019. Acessado em 21 set. 2020. Online. Disponível em: https://www.researchgate.net/publication/326506581_JVET-J1010_JVET_common_test_conditions_and_software_reference_configurations.