Miguel Oliveira
Inocêncio

Co-processador da Transformada e Quantização para AV1

AV1 Transform and Quantization Co-Processor

DOCUMENTO
PROVISÓRIO

Miguel Oliveira Inocêncio

Co-processador da Transformada e Quantização para AV1

AV1 Transform and Quantization Co-Processor

Dissertação de Mestrado apresentada à Universidade de Aveiro, para obtenção do grau de Mestre em Engenharia Eletrónica e de Telecomunicações, sob orientação do Professor Doutor António Navarro . . .

DOCUMENTO
PROVISÓRIO

o júuri / the jury

presidente / president   **ABC**
Professor Catedrático da Universidade de Aveiro (por delegação da Reitora da Universidade de Aveiro)

vogais / examiners committee   **DEF**
Professor Catedrático da Universidade de Aveiro (orientador)

                                                    GHI
Professor associado da Universidade J (co-orientador)

                                                    KLM
Professor Catedrático da Universidade N

**agradecimentos /
acknowledgements**

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum. . . .

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris. . . .

Palavras-Chave   HEVC, . . .

Resumo   Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Keywords   HEVC, . . .

Abstract   Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat
at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec non-
ummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum
massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec,
leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, sus-
cipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat
lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna.
Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam
in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus
eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus
tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit.
In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam
facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet,
enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus
eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus
quis tortor vitae risus porta vehicula.

Contents

List of Figures

List of Tables

Acronyms

AOM Alliance for Open Media.

AV1 AOM Video 1.

CMOS Complementary metal–oxide–semiconductor.

Codec Encoder-Decoder.

CRT Cathode Ray Television.

fps frames per second.

GPU Graphical Processing Unit.

HEVC High Efficiency Video Coding.

IC Integrated Circuit.

MPEG Motion Picture Experts Group.

TV Television.

UHD Ultra-High-Definition.

Glossary

Codec Encoder-Decoder. Also referred to the method of compressing and decompressing a video sequence.

Interlaced scanning Technique used by televisions for broadcasting and displaying, where only odd or even numbered lines of a frame are transmitted/displayed at a time, alternately.

libaom Reference software for AV1, released by Google in June 2018.

Progressive scanning Technique used by more recent screens, where each frame is displayed as a whole, from top to bottom, and left to right.

RGB Color space based on the addition of Red, Green and Blue components for complex color representation.

VP8/VP9 Open-format video codecs developed by Google, released in 2008 and 2013, respectively.

x

Todo list

CHAPTER 1

Introduction

1.1 | Background and Motivation

Since the spark of television research in 1887, a tremendous investment has been put into increasing the quality of images, cameras and screens that display them [1].

In the early years of mechanical television, this desire was pursued by making changes to the *Nipkow* disks, up to the decline of the mechanical TV, around the 1930's. The consequential rise of all-electronic TVs started with the capture of images with the same cathode tubes put into Cathode Ray Televisions (CRTs), with broadcasts of the live analog recordings, since there were no available methods of storing images, up to 1955, with the development of the open-reel magnetic tape [2].

The evolution of Complementary metal–oxide–semiconductor (CMOS) technologies however, led to the downfall of cathode ray tubes, and to the rise of image capture to a digital sensor, that allowed better image captures and lower demands in terms of physical storage space. However, with the desire for higher fidelity video, the quantity of information captured also increased. Whether by increasing the sensor resolution, color bit depth or frame rate, the captured video sequences have increased its size throughout the years. For instance, for a video of 640×360 (considered as a low resolution), at 30 frames per second (fps), considering each captured color (RGB) is represented with 8 bits, there is approximately 166 Million bits per second (Mbps) of captured information. This means that a short 5 minute video would occupy more than 6 Giga Bytes (GB) of memory. This aspect gets more severe once higher resolutions are considered. For newer standards such as 4K Ultra-High-Definition (UHD) (3840×2160) or 8K UHD (7680×4320), under the same conditions, a ten minute video would occupy 448 GB and 1792 GB of raw data, respectively.

Falar dos color spaces mais à frente?

To further aggravate the situation, video consumption got massively adopted on the average consumer level, and continues to grow, both in the average number of hours watched by users and in the resolutions of the video, making the bandwidth used on the visualization of video footage the highest between all other application. With the development of higher video sizes, increase of the average number of connected devices per user and overall market expansion through the number of consumers, this margin will continue to grow [3, Trends 1 & 4].

This problem has led to the introduction of a new concept: *Video Compression* [1], which is the process of reducing the size of a video sequence, while still maintaining its playback capabilities. The Codec takes advantage of redundant information present on the raw data to reduce the size of the video, without heavily modifying the original picture or its quality.

The first form of video compression, Interlaced scanning, dates from 1940, and was purely

---

[1]Also called *Video Coding.*

analog. This solution was introduced with the intent of reducing the necessary broadcasting bandwidth for old CRTs, without decreasing the displayed fps. And even though this technique has been implemented over more than seventy years, it has proven to be so efficient that most TV channels today still use interlaced broadcasting.

However, analog television is now obsolete, as well as CRTs. The massive developments in Integrated Circuit (IC) fabrication led to the rise of the digital era we now live in. Therefore, most screens (be it televisions, monitors or cellphones) use digital, Progressive scanning. As such, the use of analog compression techniques wasn't applicable. Accordingly, the evolution of digital video led to the development of digital compression techniques, such as the one presented in this work.

Being purely digital, these methodologies rely on computers and other processors to analyze data and apply the compression algorithms, making them very demanding processes from a computational standpoint. As expected, a high compression ratio is only obtainable by a high complexity algorithm, which also increases with the size of the video (more data leads to more analysis). Since in the early days of digital video, the used resolutions were lower as to the ones used in the present days, the compression algorithms used were not very demanding. However as the pursuit for higher quality video continued, so did the necessity for better compression ratios, and therefore the computational needs also increased. Such complex softwares lead to a high power consumption from the processor executing it, making such implementations unsuitable for portable, battery limited applications, such as cellphones or laptops. Besides this huge factor, such softwares tend to be very slow, specially when a real time compression or decompression is desired.

To amend for these factors, and to increase the reachability of high quality video to as many users as possible, these applications needed to have a viable solution that didn't compromise its usability. Accordingly, a new approach has been implemented on the most recent codec's. Besides the optimization of pure software compression/decompression solutions, there has been a great focus on the development of specialized hardware for such codecs. This solution could redress many of the problems presented previously, making them viable on a mobile implementation, as well as other specialized appliances, since such co-processors usually present a better performance than generic ones. This tendency has already been verified on the implementation choices on recent smartphones [4, p. 14], as well as recent *Nvidia* Graphical Processing Unit (GPU) lineups [5].

Since each compression algorithm tend to be very different from its predecessors, either by making changes to its bitstream or functioning principles, each time a new codec is released, there is a need to backup its development with a new set of hardware implementations. This makes the improvement of video compression techniques a continuous effort, in many engineering branches, as the technology needs to keep up with the demands of consumers, in a variety of applications.

Due to the broad access to video, and its influence in a variety of markets (besides video consumption itself), big companies have made investments on the improvement of video quality, and respective compression algorithms. These investments have provoked somewhat of a "Codec War". Since 2010, several video codecs have been deployed, and quickly replaced by a newer version, which presents better compression gains, at a lower quality degradation, such as the replace of *VP8* (released in 2008) with *VP9* (2013).

1.2 | Scope

*Complexity analysis and necessity for hardware solutions

*Thesis objectives

AOM Video 1 (AV1) is the most recently released[2] video codec. It was developed as a Joint Development Foundation [6] project, under the name of Alliance for Open Media (AOM). This codec took the same objective as its main predecessor, *VP9*, which was to be an open source, royalty free alternative to Motion Picture Experts Group (MPEG)'s state of the art video codec, *High Efficiency Video Coding (HEVC)*.

Upon release, *VP9* rivaled *HEVC's* performance. However, soon after, the market demanded higher compression performance, giving origin to consortium of enterprises that now represent AOM, and to the development of *AV1*, in 2015. The first release of this coding format was made in March 2018, with the first release of its reference software, *libaom*, being made three months later, in June 2018.

Besides its main objectives, *AV1* was also developed with the intent of being implementable in hardware. Therefore, various design choices were made to make the algorithm low memory consuming, and highly parallelizable.

The desired compression performance was obtained at the cost of a highly complex algorithm (and reference software), that severely outperforms *VP9*, at the cost of much higher compression times [7] .

referência para secção mais à frente?

Taken these factors, there is a high demand for dedicated hardware architectures, that can speed up the compression/decompression times and reach real-time usability on live-streaming applications, such as video-conferencing, live-content visualization, etc.

1.3 | Outline

*General outline of the different chapters **Last section to do**

References

[1]   Mark Schubin. "What Sparked Video Research in 1877? The Overlooked Role of the Siemens Artificial Eye [Scanning Our Past]". In: *Proceedings of the IEEE* 105.3 (Mar. 2017), pp. 568–576. ISSN: 0018-9219, 1558-2256. DOI: 10.1109/JPROC.2017.2652998.

[2]   Marco Jacobs and Jonah Probell. "A Brief History of Video Coding". en. In: (), p. 6.

[3]   *Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White Paper.* en. https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html.

[4]   Scientiamobile. *Mobile Overview Report April – June 2018.*

---

[2]Currently, there are other codecs being developed, without official bitstream release

[5]   *Video Encode and Decode GPU Support Matrix.* en. https://developer.nvidia.com/video-encode-decode-gpu-support-matrix. Nov. 2016.

[6]   *Joint Development Foundation.* en. http://www.jointdevelopment.org/.

[7]   Dan Grois, Tung Nguyen, and Detlev Marpe. "Performance Comparison of AV1, JEM, VP9, and HEVC Encoders". In: *Applications of Digital Image Processing XL*. Ed. by Andrew G. Tescher. San Diego, United States: SPIE, Feb. 2018, p. 120. ISBN: 978-1-5106-1249-5 978-1-5106-1250-1. DOI: 10.1117/12.2283428.

CHAPTER 2

Basics of Video Compression

2.1 | Human Visual System

*Essency of video compression relies on making changes the image without serious perception by the user

*Eye Functioning

*"Known Issues" (lower perception to chroma, high frequencies, etc)

*Opportunity to explore various types of redundancies to the image

2.2 | Redundancy Exploitation

*Types of redundancies (Temporal, Statistical and Coding)

*Color subsampling

*Intra-prediction

*Inter-prediction

*Transform and Quantization

*Entropy Coding

2.3 | Basic Video Compression/Decompression System

*Encoder Model

*Decoder Model

CHAPTER 3

AV1 Video Codec

3.1 | History

*Previous generations

*Development Process

*AOMedia companies

3.2 | Terms and Definitions/Terminology

Review *AV1 Bitstream and Decoding Process*

3.3 | Key Aspects and Innovations

*Comparation with past generations

*Introduction of modules not present on other video codecs

*Block diagram

3.4 | Performance Analysis

*Compression gains

*Quality assessment

*Complexity (general/modules) and timing issues

CHAPTER 4

Developed Architecture

4.1 | REEEEEEEEEEEEEEEEEEEEEEEEE

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.