University of Warsaw
Faculty of Mathematics, Informatics and Mechanics

Adam Deryło
Student no. 342007

Adrian Hess
Student no. 342013

Magdalena Pałkus
Student no. 231023

Michał Skwarek
Student no. 777321

GPU acceleration of CCSDS Rice decoding

Bachelor's thesis
in COMPUTER SCIENCE

Supervisor:
Paweł Gora
Faculty of Mathematics, Informatics and Mechanics, University of Warsaw

Warsaw, Jan 2023

Abstract

TBD

Keywords

CCSDS Rice coding, GPU, Compute Unified Device Architecture (CUDA)

Thesis domain (Socrates-Erasmus subject area codes)

11.3 Informatics, Computer Science

Subject classification

D. Software
D.1.3. Concurrent Programming
I.4.2. Compression (Coding)

Tytuł pracy w języku polskim

Akcerleracja GPU dekodowania CCSDS Rice

Contents

Introduction 5

1. Key terms 7
1.1. FITS 7
1.2. RICE 8

2. GPU acceleration of CCSDS Rice decoding algorithm 9
2.1. Naive approach 9

Bibliography 11

Introduction

Data compression is a widely utilized technique for reducing the storage requirements and transmission time for large data sets. However, when it comes to training deep learning models, the decompression of compressed data can create a bottleneck in the machine learning pipeline, particularly when dealing with specialized data formats such as astronomy and medicine that employ custom compression algorithms that can be computationally expensive to decompress.One such specialized compression algorithm is RICE coding, which is widely used in the FITS data format in astronomy. To address this bottleneck, utilizing Graphics Processing Units (GPUs) and parallelization techniques have emerged as promising solutions for accelerating the decompression of large data sets by leveraging the parallel processing capabilities of GPUs. While established solutions for mainstream lossless compression algorithms like JPEG-2000 exist, this paper aims to investigate the potential of GPU acceleration and parallelization in enhancing the performance of RICE coding, a specialized and niche compression algorithm.

Chapter 1

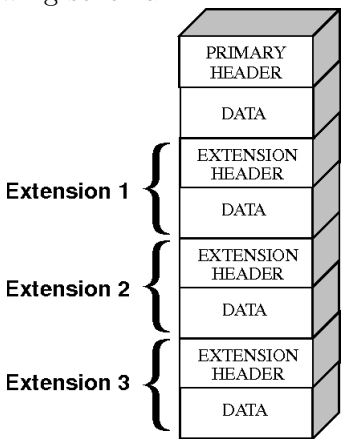
Key terms

1.1. FITS

The Flexible Image Transport System (FITS) format is a digitalfile format that is commonly utilized in the field of astronomy for storing and transferring scientific data. FITS files typically contain images, data cubes, and tables of observational data, as well as associated metadata. One of the key characteristics of FITS is its ability to store multiple data arrays in a single file, which allows for the efficient storage and transfer of large data sets. A FITS files are composed of following FITS structures, in the order listed:

- Primary header and data unit (HDU).
- Conforming Extensions (optional).
- Other special records (optional, restricted).

Thus, files usually resamble following schema:



Where, all headers including the primary one contain relevant metadata as a list of keys and value pairs. Furthermore, according to most recent FITS standard published by NASA, there are three types of standard extensions:

- IMAGE extensions.
- TABLE ASCII-table extensions; and
- BINTABLE binary-table extensions

1.2. RICE

Chapter 2

GPU acceleration of CCSDS Rice decoding algorithm

2.1. Naive approach

Bibliography

[Bea65] Juliusz Beaman, *Morbidity of the Jolly function*, Mathematica Absurdica, 117 (1965) 338–9.