Computer Vision Group Project Proposal

— Context-based Trit-Plane Coding for Progressive Image Compression

Group members: Jiajun Li, Hanliang Jiang, Haoxiang Li

Project title: Reproduction and Optimization of Context-based Trit-Plane Coding Algorithm

Project description and goals

As a crucial problem in image processing, image compression has been developing for decades and many well-known codes, such as JPEG and JPEG2000, have been made. However, even though these codecs generate decent rate-distortion (RD) results, these images still fail to catch up with the growing need for better-quality compressed images. Therefore, many Deep Learning methods have been made by scientists in search of higher RD performance in image compression.

Our group found a paper dedicated to using a context-based trit-plane coding image compression algorithm for better RD results and decided to reimplement it and make an optimization based on the original models.

In order to understand and fine-tune this algorithm, we decided to read this paper and other related works in-depth and reimplement it using the CLIC dataset, JPEG-AI dataset and Kodak dataset. Following that, we will try to understand each module's function, preparing for the potential optimization. Subsequently, we would like to make an attempt to fine-tune this algorithm and test it on the aforementioned dataset. We will then evaluate the RD performance of the optimized algorithm with this paper’s algorithm as well as other related algorithms, if possible.

Member roles

Three team members worked together on every stage of the project. We have already collected datasets from the Internet. Furthermore, we will follow the outline of each stage and schedule team meetings periodically, exchanging thoughts and discussing solutions. Since we didn’t have experience in image compression, we decided to accomplish our project objectives collectively. Though fine-tuning tasks and evaluation tasks may be processed separately for efficiency, most of our work will remain to be done as a team.

Resources

The dataset of our project will be the CLIC dataset, JPEG-AI dataset and Kodak dataset. These datasets contain multiple images for image compression training and benchmarking. The CLIC dataset and JPEG-AI dataset typically have image sizes ranging from 2 to 5 megabytes (MB) per image, while the Kodak dataset has a relatively smaller image size, less than 1 megabyte (MB).

Figure 1. Example of CLIC dataset

Figure 2. Example of JPEG-AI dataset

Figure 3. Example of Kodak dataset

Our implementation is based on Python 3.9, with several machine learning libraries including pytorch, torchvision as well as CompressAI.

We plan to make some changes in the framework of the original algorithm and there is a possibility that we will use outside codes which have proven to be efficient in increasing RD performance. We plan to introduce some novel convolution layers if we find high-performance methods by that.

We haven't got access to Machine Learning cloud services yet, but we have looked into multiple services by Amazon AWS, Google Cloud and Microsoft Azure. We may choose Google Cloud's Vertex AI Platform as the platform to perform our intensive computer vision jobs.

Reservations

It is possible that the initial setup and configuration of the project may be difficult since the project will be run on different environments, requiring precise control of software tool versions. Another obstacle may be the algorithm, where little changes can be made to achieve even better results than the originally proposed method. Still, the most realistic goal that is supposed to be accomplished is to run the compression algorithm successfully and test its effectiveness as described in the paper.

Relationship to our background

All three team members are majoring in electrical and computer engineering, and do not possess **any** prior experience in computer vision. Each member is capable of writing basic python code and utilizing popular packages like numpy or pandas.