Abstract

Block partitioning in video encoding with the use of artificial neural network

This dissertation explores the field of video encoding, where moving images are compressed for to be efficiently transmitted in television systems, Video-On-Demand platforms, and similar applications. The research presented in this dissertation is focused on designing video encoder control algorithms. Special emphasis is put on CTU (Coding Tree Unit) partitioning, the most computationally intensive part of the encoding process. The research aims to develop a partitioning algorithm that significantly reduces the computational complexity of the encoder while preserving the compression efficiency, compared to the existing solution in reference encoder. Thus, the research explores Artificial Neural Network (ANN)-based approaches for partitioning algorithms.

The dissertation begins with a description of CTU partitioning in HEVC, along with the contextuality of the decisions. A survey of the existing partitioning algorithm is presented, pinpointing the aspects that require improvements. Methods that offer control over the trade-off between coding time and coding efficiency are reviewed, pointing out the complexity of controlling this trade-off in existing solutions.

Two approaches to the partitioning problem are explored: one for decision-making at Coding Unit (CU) scope and one for joint Coding Unit (CU) scope and Prediction Unit (PU) scope. ANN with non-trivial decision algorithms is introduced. Proposed ANNs are designed to jointly estimate depth-level probabilities for individual CTU subareas. A custom training dataset has been prepared to train the ANN models. The ANN architecture development process is described, and a detailed training and evaluation results analysis of the final models is presented.

This dissertation proposes original, non-trivial decision algorithms that utilize probabilities determined by the ANN. These decision algorithms are defined in two variants: hard-decisive and soft-decisive. It is demonstrated that one of the proposed algorithms allows a straightforward control of the trade-off between coding time and coding efficiency through a single parameter.

Comparative analysis with state-of-the-art solutions demonstrates that the proposed partitioning algorithms offer the best trade-off between encoding time reduction and coding efficiency.

Considering control over the trade-off between encoding time and coding efficiency, the proposed method provides the best results and the most straightforward control among other state-of-the-art solutions. Additionally, a new metric for rapid comparison of such methods is introduced, which coincides with well-established evaluation approaches.

Lastly, this dissertation explores the impact of contextual effects on partitioning decisions. The key achievements are:

- The author's experiment on determining the impact of the encoding context on decisions.
- Modifications to the proposed ANN model for processing contextual information.
- The author's method of training an ANN with augmented ground truth data.
- The author's method for global partitioning patterns optimization which introduces a negligible increase in encoding time.

All experiments were conducted using a modified version of HEVC reference model software developed by the author. This modification, which enables the rapid implementation of ANN-based partitioning algorithms, has been released under open-access terms.