ARIA
APPLIED RESEARCH IN ACTION

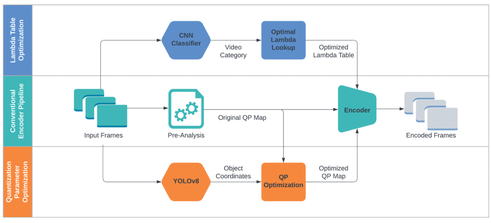
Content-Aware Video Encoder Optimization

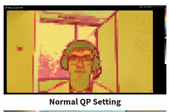
Elevating conventional video encoder with deep learning models for enhanced video quality and bitrate efficiency

Alex Xu

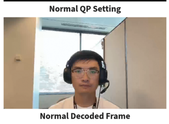
Qiang Sun
ACADEMIC SUPERVISOR

Jeff Jiang and Yu Shao
INDUSTRY SUPERVISORS




Normal QP Setting


Improved QP Setting


Normal Decoded Frame


Improved Decoded Frame

PROJECT SUMMARY

Modern video codecs leverage both spatial and temporal redundancy to compress raw video data into compact bitstreams for efficient storage and transmission. However, these conventional compression techniques operate unselectively with respect to video content, often leading to suboptimal compression outcomes in specific applications. For instance, empirical investigations [1] reveal that the default rate-distortion lambda tables employed in codecs such as AVC and HEVC can be further optimized based on the inherent characteristics of the video content, thus yielding improvements in the overall BD-rate. In response to this problem, we propose a content-adaptive methodology aimed at enhancing the encoding efficiency of video codecs. Our approach addresses two key aspects:

(1) Optimal Lambda Table Tuning: To tailor the encoding process to different video categories (e.g., gaming, nature, and screen content), we integrate a convolutional neural network capable of classifying the video into different categories. This information serves as a valuable cue for selecting the optimal lambda table, resulting in a noteworthy average PSNR-BD-rate reduction of 2.6% for gaming videos and 2.8% for nature-themed content.

(2) Object Detection Based Quantization Parameter (QP) Optimization: In scenarios where foreground objects are more critical than the background (e.g., video conferencing focusing on people), we integrate YOLOv8 [2] to track the object position in real time. By dynamically re-weighting QP to prioritize the region of interest (foreground), we achieve a notable average of 3.6% PSNR improvement in the quality of foreground objects while maintaining the overall BD-rate.

REFERENCES

[1] F. Zhang and D. R. Bull, "Rate-Distortion Optimization Using Adaptive Lagrange Multipliers," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 10, pp. 3121-3131, Oct. 2019, doi: 10.1109/TCSVT.2018.2873837.
[2] G. Jocher, A. Chaurasia and J. Qiu, YOLO by Ultralytics, https://github.com/ultralytics/ultralytics, 2023.

AMD

Computer Science
UNIVERSITY OF TORONTO
Master of Science in Applied Computing