MACHINE LEARNING BASED SYMBOL PROBABILITY DISTRIBUTION PREDICTION FOR ENTROPY CODING IN AV1

Kshitiz Kumar
AI22MTECH02002

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Department of Artificial Intelligence

April 25, 2022

Problem Statement

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- Overcoming the deficiencies of lookup table based scheme in case of entropy encoding, to achieve more accurate prediction of symbol probability distribution.

Motivation

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- The symbol probability is derived with handcrafted context models and lookup tables that store the predicted probabilities corresponding to different entropy contexts.
- But there are some fundamental deficiencies in the lookup table based scheme:

    1) The entropy context features have to be discrete so that they can be used to index the lookup tables.
    2) To reduce the size of the lookup table, the number of contexts cannot be very large.
    3) The probability distributions stored in the lookup tables are maintained separately without taking their correlations into consideration.

Entropy Coding

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- Entropy coding (or, encoding) is a lossless data compression scheme to compress digital data by representing frequently occurring patterns with few bits and rarely occurring patterns with many bits.
- It is done on an image after the quantization stage. It enables to represent an image in a more efficient way with smallest memory for storage or transmission.
- In most images their neighboring pixels are correlated and thus contain redundant information. We want to find less correlated representation of the image, then perform redundancy reduction and irrelavancy reduction.
- Redundancy reduction removes duplication from the signal source (for instance a digital image).
- Irrelavancy reduction omits parts of the signal that will not be noticed by the Human Visual System (HVS).

Entropy Coding (contd.)

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- Entropy encoding further compresses the quantized values in lossless manner which gives better compression in overall.
- It uses a model to accurately determine the probabilities for each quantized value and produces an appropriate code based on these probabilities so that the resultant output code stream will be smaller than the input stream.
- It is true that the quantization does compress the data tremendously. After the quantization step many of the pixel values are either eliminated or replaced with other suitable values. However, those pixel values are still represented with either 8 or 16 bits.
- Aim is to minimize the number of bits used, by means of entropy encoding.
- Examples of Entropy Coding Schemes: Kahunen-Loeve transform (or, PCA), Kolmogorov Entropy, Huffman Coding, Arithmetic Coding, etc.

Arithmetic Coding

- One of the latest and most popular encoding scheme.
- It's an entropy-based algorithm, first proposed in a paper from 1987 (Witten, Ian H., Radford M. Neal, and John G. Cleary. "Arithmetic coding for data compression." *Communications of the ACM* 30.6 (1987): 520-540).
- AE encodes the entire message using a single number between 0.0 and 1.0 .
- Each symbol in the message takes a sub-interval in the 0-1 interval, corresponding to its probability.
- To calculate the probability of each symbol, a frequency table should be given as an input to the algorithm. This table maps each character to its frequency.
- The more frequent the symbol, the lower number of bits it is assigned. As a result, the total number of bits representing the entire message is reduced.

Arithmetic Coding (contd.)

- In contrast to AE, Huffman coding uses a fixed number of bits for all symbols in message so it is less efficient than AE.

Video Coding

- Process of video encoding is dicatated by “Video Codecs” or “Video Compression Standards”.
- CODEC : enCOder and DECoder
- Each codec is comprised of –

1) An Encoder: to compress the video
2) A Decoder: to recreate an approximation of the video

- Examples of Video Codecs: H.264, VP8, VP9, RVP40, AV1, etc.

AV1 Video Codec

- AV1 is a state-of-the-art open-source and royalty-free video compression format that is jointly developed and finalized in early 2018 by the Alliance for Open Media(AOM) industry consortium.
- It is owned by Google and has succeeded its predecessor VP9 to compete with HEVC/H.265 .
- Netflix is right now using this for streaming its videos.
- In AV1, multi-symbol arithmetic coding is adopted.
- It achieves significant bit-rate savings over previous codecs such as VP9 and HEVC.
- For intra prediction, AV1 supports up to 56 directional modes, DC prediction mode, path prediction mode, quadratic interpolation filtering modes, the palette mode, the intra block-copy mode, as well as a SIMD friendly variation of the recursive-filtering-based modes.

Entropy Coding in AV1 codec (Encoder Part)

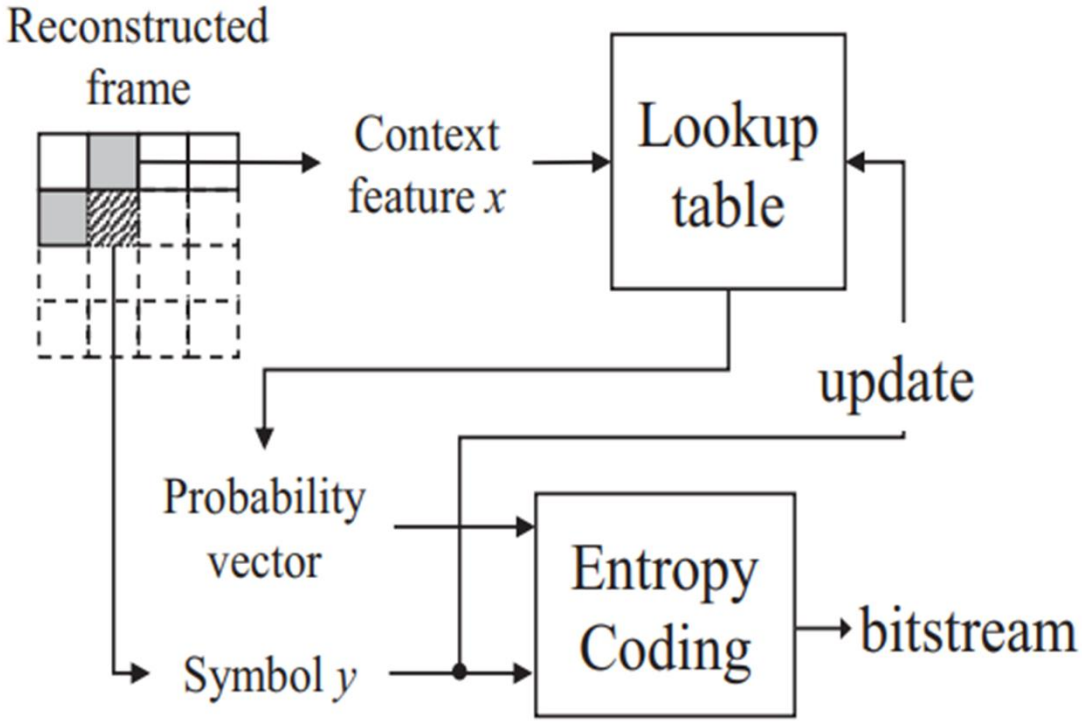

Fig.1: The shaded block in the reconstruction frame denotes the current block; the grey blocks denote the adjacent blocks where the entropy context features can be extracted.

Proposed Model

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- In the proposed approach, an ML model G is being utilized to predict the probability distribution of a given symbol/signal.
- To encode a symbol in the current block, the context features are first extracted from the adjacent reconstructed blocks, and then fed into G to calculate the predicted probability distribution.
- According to the computed probability prediction, the symbol is encoded into the bit-stream via entropy coding.
- At the decoder side, the ML model G plays an identical role as it does in the encoder to provide symbol probability distribution prediction.
- Then, entropy decoding is applied to decode the symbols from the bit-stream.

Proposed Model (contd.)

- The models for the encoder and decoder share the same initial weights and update rate to guarantee the synchronization of the encoding and decoding processes.
- The initial weights of the models are pre-trained with the statistics collected by running the AV1 encoder on some representative training video sequences.

Proposed Model (contd.)

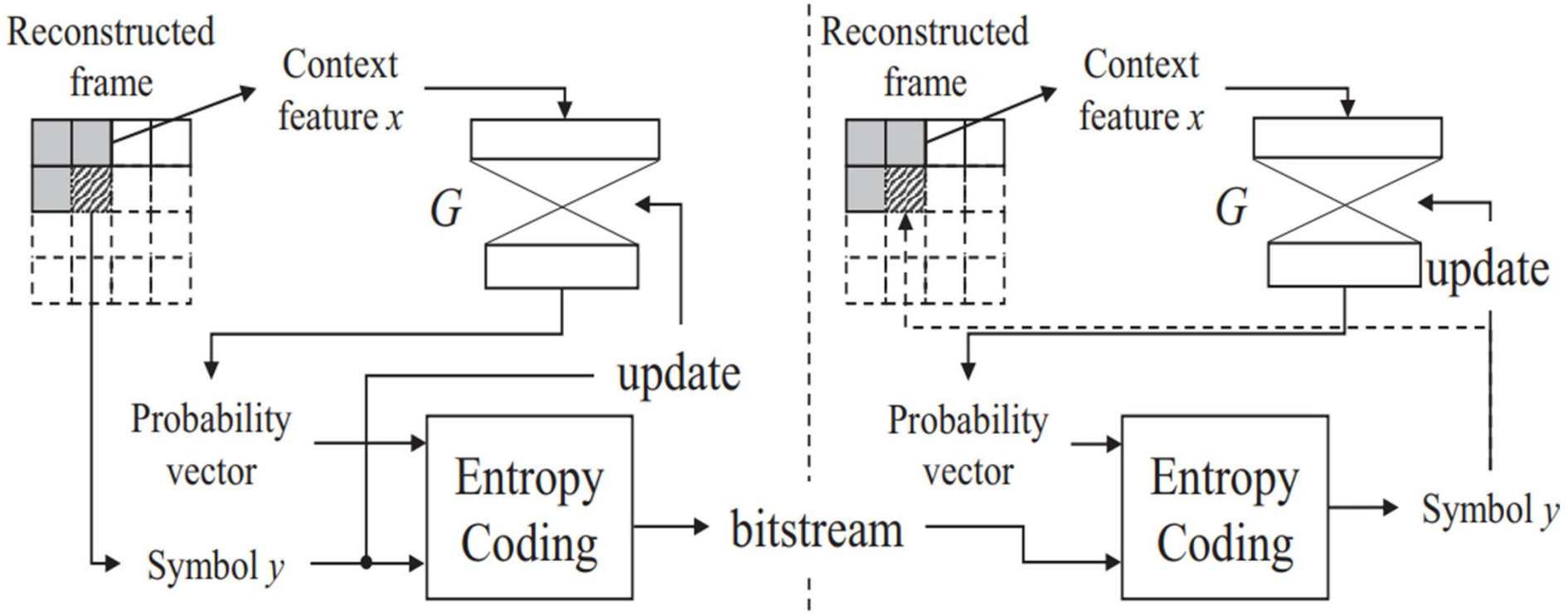

Fig.2 : Entropy coding pipeline

Performance on AV1

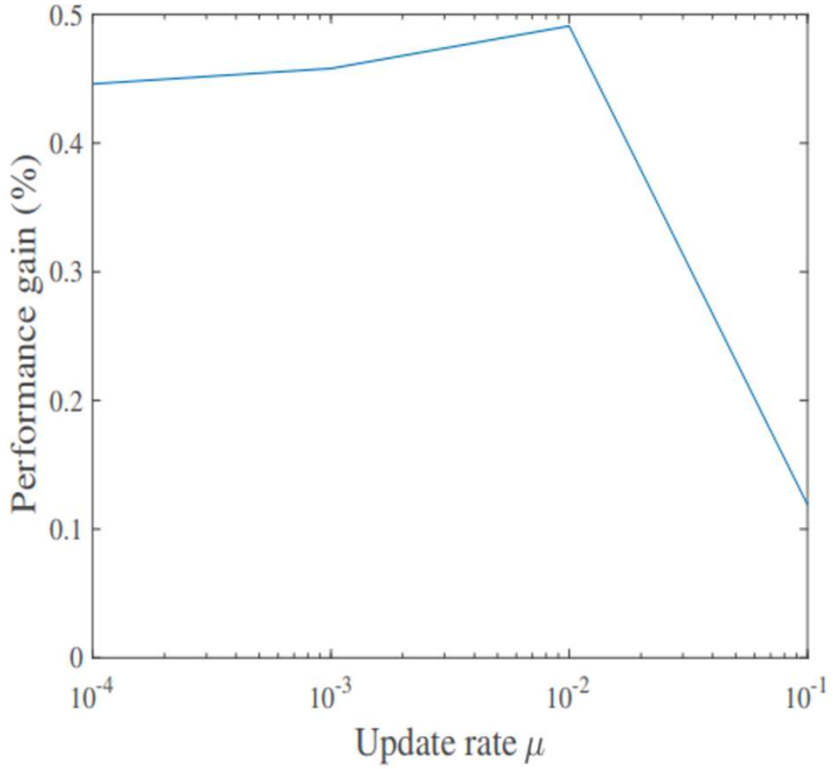

Fig.3 : PSNR BD-rate gains with different update rates.

Performance on AV1 (contd.)

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

		Low resolution	Medium resolution	High resolution
Y	W/O update	0.472%	0.445%	0.413%
	W/ update	0.495%	0.493%	0.503%
Y+UV	W/O +update	0.628%	0.512%	0.455%
	W/ update	0.698%	0.672%	0.686%

Table 1. Performance gains of the proposed model

Conclusion

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- ML models can achieve more accurate probability prediction than the lookup table based scheme used in state-of-the-art codecs, leading to better compression performance.
- The proposed approach was implemented and tested on AV1 for the entropy coding of intra prediction modes.
- Experimental results demonstrate that it can improve the efficiency of entropy coding significantly.

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

THANK YOU