Entropy Coding of Neural Network Latent Space Coefficients for Point Clouds

Dominik Mehlem, Mirco Dilly

Institut für Nachrichtentechnik, RWTH Aachen University

Point Cloud Compression

Point Clouds

- 3D collection of points representing objects
- Geometry: points in space $(x, y, z) \in \mathbb{R}^3$
- Attributes: color information, normal vectors, ...

Compression Necessary

- Raw point cloud with approx. $0.7 \cdot 10^6$ points per frame and 30 fps
- 10 bit geometry precision, 8 bit attribute precision
- Every point requires $3 \cdot (10 + 8)$ bit $= 54$ bit
- Due to data types up to 24 Byte per point:

$$R = 0.7 \cdot 10^6 \, \frac{\text{points}}{\text{frame}} \cdot 30 \, \frac{\text{frames}}{\text{s}} \cdot 24 \, \frac{\text{Byte}}{\text{point}} \approx 500 \, \frac{\text{MByte}}{\text{s}}$$

Autoencoder-based Geometry Compression [1]

Architechture

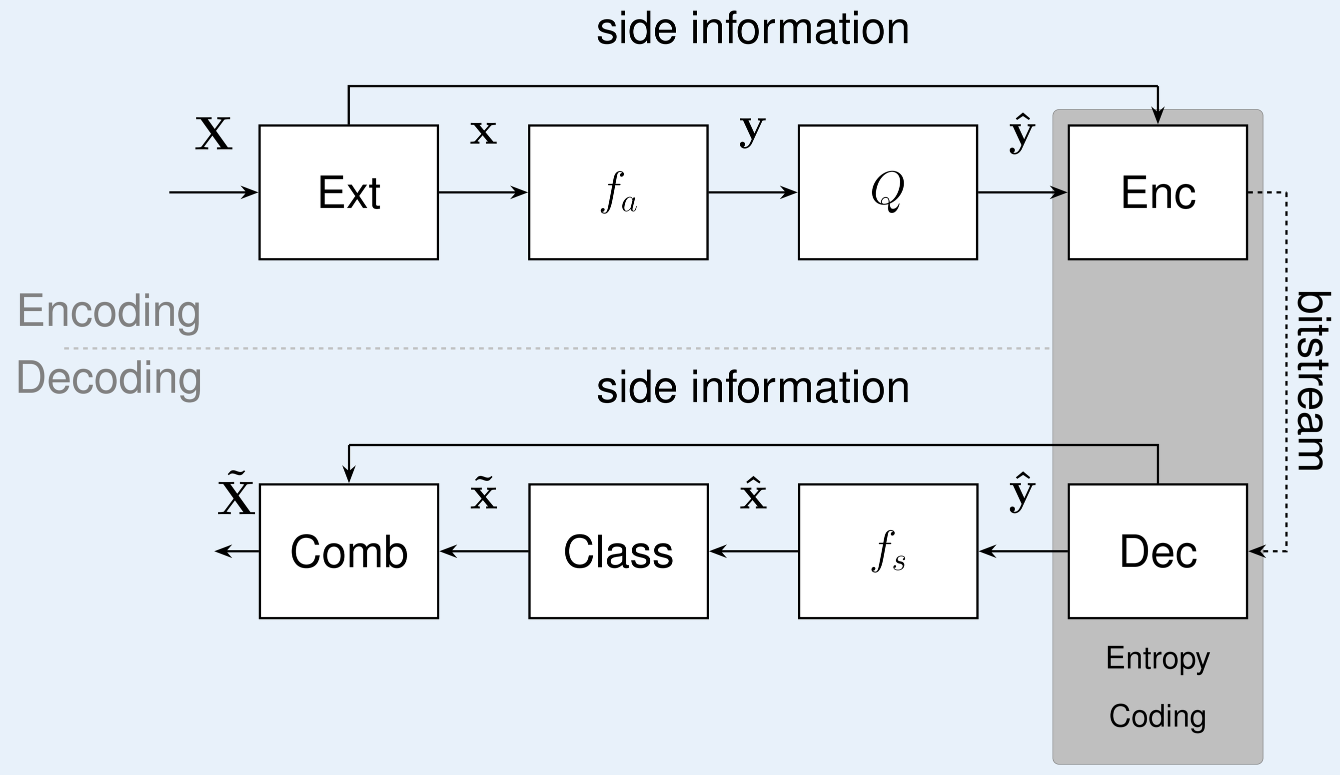
- Static geometry compression
- Size of block tensor (input) $x : 64 \times 64 \times 64$
- Size of LSC tensor (output) $y : 8 \times 8 \times 8 \times 32$

Training

- Loss function: Rate and focal loss $\mathcal{L} = R + \lambda \sum_{z \in X} \mathrm{FL}(z)$

Compression Ratio

$$C = 1 - \frac{r_{\text{enc}}}{r_{\text{ref}}}$$



Entropy Coding

Coding of Geometry LSCs

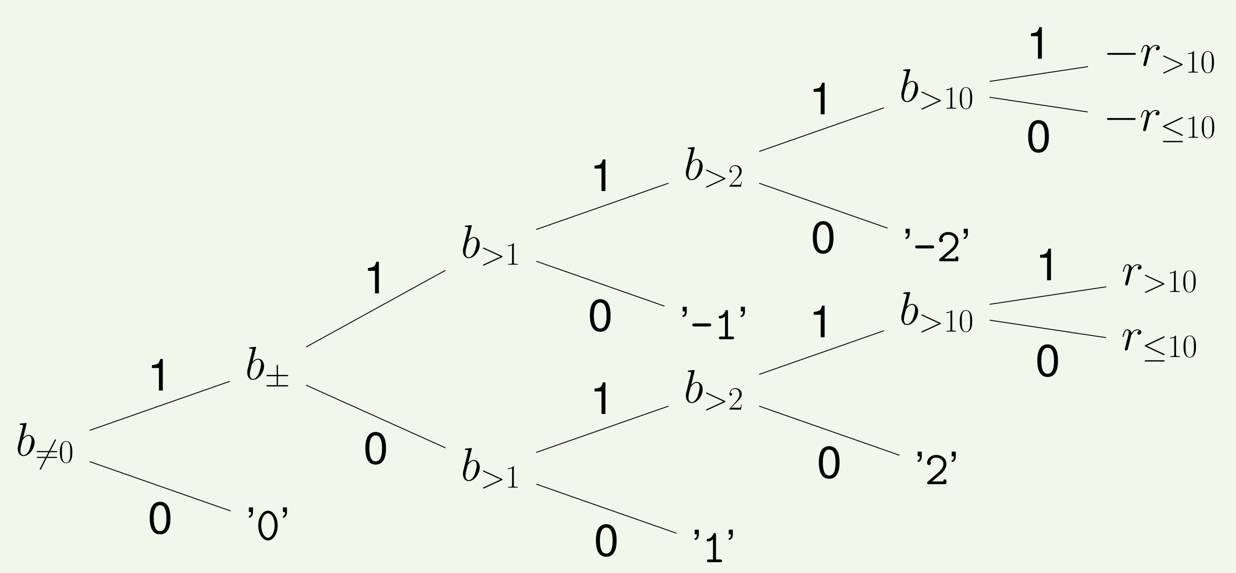
- Lossless coding
- Input: Quantized latent space coefficients

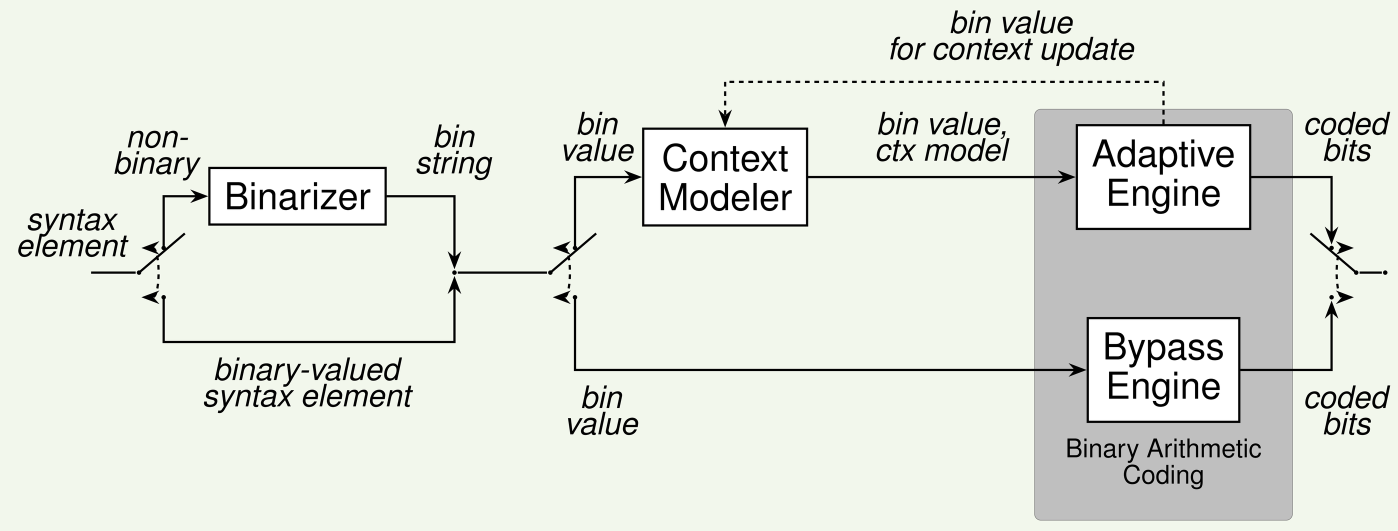
Different Methods tested

- Huffman Coding (HC)
- Adaptive Huffman Coding (AHC)
- Arithmetic Coding (AC)
- Context-based Adaptive Binary Arithmetic Coding (CABAC)
- Dictionary-based Coding
  - LZMA2 (7zip)
  - Deflate (Gzip)

$\Rightarrow$ Dictionary-based methods and CABAC show most promising results

CABAC – Block Diagram and Binarization [3]





Binary Decision Bits

- $b_{\neq 0}$: Significance bit
- $b_{\pm}$: Sign bit
- $b_{>1}$: Greater-1 bit
- $b_{>2}$: Greater-2 bit
- $b_{>10}$: Greater-10 bit

Remainder Coding

- $-r_{\leq 10}$: Fixed-length binarization
- $-r_{>10}$: Exponential Golomb coding

CABAC – Context Modeling

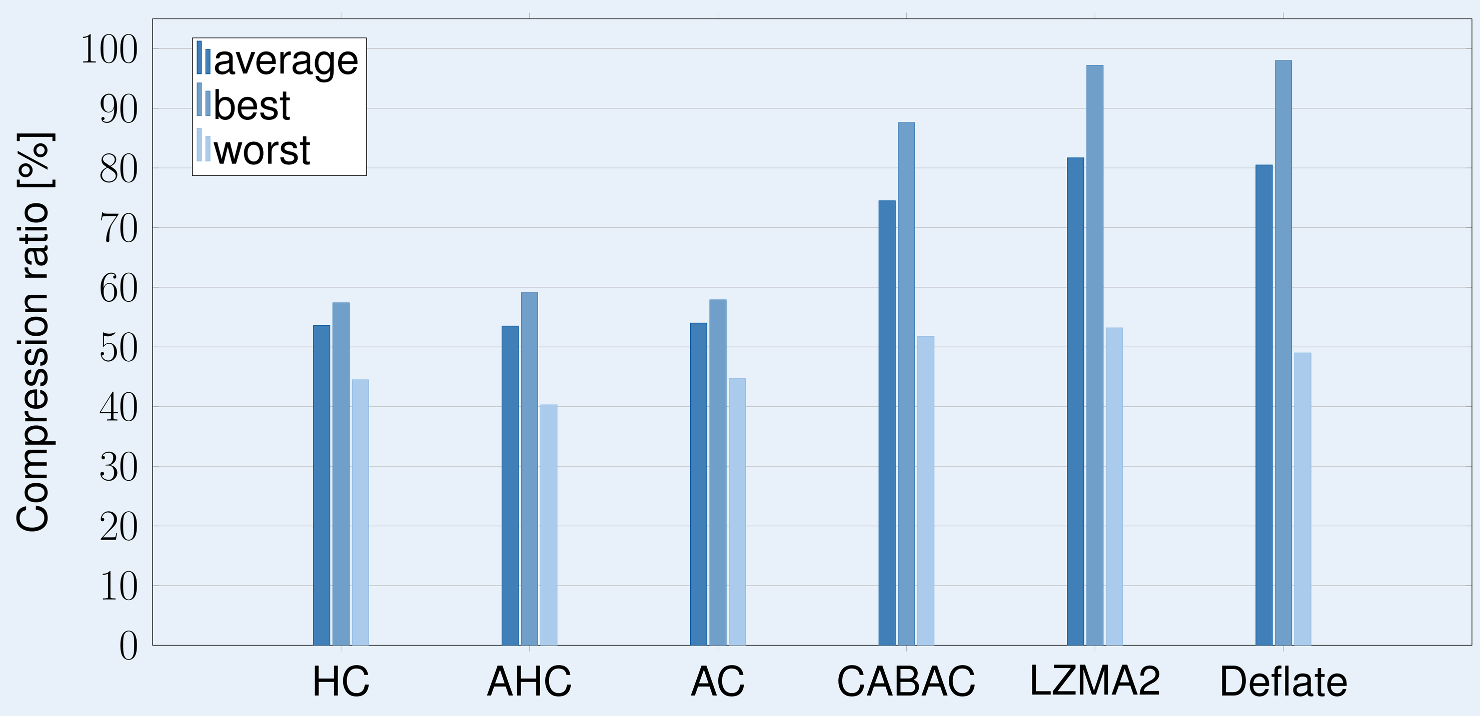
35 Contexts

- Sign bit and remainder >10 bypass coded
- Significance bit, Comparison bits with third-order model
- Remainder <10 with zeroth-order model (only marginal probability)

Bins	Binarization	CABAC Engine	Model Order	Context ID	Template
$b_{\neq 0}$	Coding Tree	Adaptive	3rd	0-7	$\boldsymbol{T}_3$
$b_{\pm}$	Coding Tree	Bypass	–	–	–
$b_{>1}$	Coding Tree	Adaptive	3rd	8-15	$\boldsymbol{T}_3$
$b_{>2}$	Coding Tree	Adaptive	3rd	16-23	$\boldsymbol{T}_3$
$b_{>10}$	Coding Tree	Adaptive	3rd	24-31	$\boldsymbol{T}_3$
$b_{r_{\leq 10},1}$	Fixed-length	Adaptive	0th	32	–
$b_{r_{\leq 10},2}$	Fixed-length	Adaptive	0th	33	–
$b_{r_{\leq 10},3}$	Fixed-length	Adaptive	0th	34	–
$b_{r_{>10},i}$	Exp-Golomb	Bypass	–	–	–

Results

- AHC and HC with density estimation show similar results
- CABAC: Conditional dependencies within LSCs can be exploited
- Dictionary-based coding achieves highest compression



References

[1] Maurice Quach, Giuseppe Valenzise, and Frederic Dufaux. "Learning Convolutional Transforms for Lossy Point Cloud Geometry Compression". In: 2019 IEEE International Conference on Image Processing (ICIP) (Sept. 2019)

[2] Eugene d'Eon, Bob Harrison, Taos Myers, and Philip A. Chou. 8i Voxelized Full Bodies - A Voxelized Point Cloud Dataset. ISO/IEC JTC1/SC29 Joint WG11/WG1 (MPEG/JPEG) input document WG11M40059/WG1M74006. Jan. 2017.

[3] D. Marpe, H. Schwarz, and T. Wiegand. "Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard". In: IEEE Transactions on Circuits and Systems for Video Technology 13.7 (2003), pages 620-636.3

mehlem@ient.rwth-aachen.de

www.ient.rwth-aachen.de

Institut für Nachrichtentechnik, Melatener Str. 23, 52074 Aachen

SVCP 2022