

潜在変数モデルと Bits-Back 符号化を用いた画像圧縮方式における 画像の分割方式に関する検討

杉山 遥 鈴木 海友 松田 一朗
東京理科大学創域理工学部

1. はじめに

近年、潜在変数を用いた画像生成モデルと bits-back 符号化を組み合わせた可逆画像符号化手法[1]が提案されており、特に拡散モデル(Diffusion model)のような推論精度の優れた生成モデルとの組み合わせで高い符号化性能を達成する可能性が示されている[2]. 筆者らは先に、大量の画像をまとめて符号化する際に、bits-back 符号化で要求される初期ビット列という付加情報の生成法を変更することで実用的な符号化効率を達成できることを示した[3]. 本稿では、符号化対象画像を一般的な解像度の画像に拡張するために、画像の分割方式が符号化効率に与える影響を考察する。

2. bits-back 符号化と拡散モデル

bits-back 符号化は、Last-In First-Out(LIFO)バッファと同様に、ビット列の追加(push), 取り出し(pop)機構を備えるエントロピー符号化アルゴリズムである. このアルゴリズムは, pop 操作と負の情報量を対応させることで, 潜在変数モデルを用いた可逆符号化における符号化効率を改善している. 図1に bits-back 符号化手順および復号手順の基本的概念を示す. まず符号化開始時に画像 x とは別に十分な量の初期ビット列を用意しておき, pop 操作によって T 段の潜在変数 $z_i (i = 1, 2, \dots, T)$ すべてを復号(サンプリング)する. 次に x を push 操作によって符号化し, さらに全ての潜在変数 z_i を符号化(push)する. この手順において, 前半の pop 操作は後半の潜在変数の符号化(push)に伴う情報量の増加をキャンセルしていると解釈できる[1]

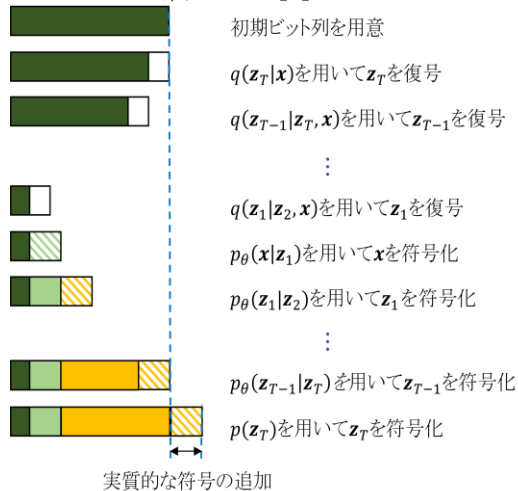


図1. 拡散モデルを用いた bits-back 符号化の手順

3. 符号化対象画像の分割

bits-back 符号化と拡散モデルを組み合わせたコーデックでは, 符号化対象画像の大きさがモデルの入力サイズによって制約を受ける. [3]では, 拡散モデルの学習・評価ともに CIFAR-10 データセット(32×32 画素)を用いていたため, コーデックの入力サイズもこれに準じている. モデルの入力サイズよりも大きい画像で符号化を行うには, 符号化対象画像を分割してコーデックに入力する必要がある.

本稿ではこの符号化処理の枠組みにおいて画像の分割方式が符号化効率に与える影響を調査する.

4. 実験結果

表1. 符号化レートの比較 (bits/pel)

bits-back		JPEG-LS
ブロック分割のみ	ライン分割 +ブロック分割	分割なし
11.01	10.85	12.59

Improved DDPM($T=100$)と bits-back 符号化を組み合わせた画像の可逆符号化方式を実装し, ImageNet の学習用画像を 64×64 画素にリサイズしたものを学習に用いた. 評価実験では符号化対象画像に Kodak データセットの画像を 512×512 にクロップしたものをを用いた. 表1に対象画像をすべて 32×32 のブロックに分割した場合, 初期ビットの生成用に JPEG-LS に入力する画像は幅が32, 高さが512のライン状に分割し, bits-back 符号化では 32×32 のブロックに分割した場合の実験結果を示す. JPEG-LS に入力する画像サイズが大きいほうが, わずかだが符号化効率が改善されることがわかる.

参考文献

- [1] J. Townsend, "Lossless compression with latent variable models," Ph.D. dissertation, Dept. of Computer Science, Univ. College London, <https://arxiv.org/abs/2104.10544arXiv:2104.10544> [cs.LG], Apr. 2021.
- [2] J. Berner, "Bits-back coding with diffusion models," https://github.com/facebookresearch/NeuralCompression/tree/main/projects/bits_back_diffusion
- [3] H. Sugiyama, *et al.*, "Initial bits generation for lossless image compression using bits-back coding with a latent variable model," Proc. 2024 International Workshop on Advanced Image Technology (IWAIT 2024) (to appear).