

レートをもっとする予測器と算術符号を用いた画像の可逆符号化

Lossless Image Coding Using Minimum-Rate Predictors and Arithmetic Code

白井 規之*

松田 一朗*

伊東 晋*

Noriyuki SHIRAI

Ichiro MATSUDA

Susumu ITOH

1. はじめに

我々は先に、レートをもっとする予測器を用いた画像データの可逆符号化方式を提案した [1]。この方式では、コンテキストモデリングに基づいた適応的なエントロピー符号化を前提として予測誤差の情報量を定式化し、その値が最小となるように予測器を最適化することで高い符号化効率を実現している。しかし、エントロピー符号化部に画素単位の可変長符号を採用していたため、情報量が 1bit/pel 以下となるような単純な画像において符号化効率が大きく低下する場合があった。本稿では、可変長符号の代わりに多値算術符号化方式の一種であるレンジコーダ [2] を導入し、上記問題の解決を図る。

2. レートをもっとする予測器の設計 [1]

画素 p_k の輝度値を $S(p_k)$ と表記したとき、注目画素 p_0 に関して線形予測を実行した際の予測誤差信号 e は次式で表せる。

$$e = S(p_0) - \sum_{k=1}^K a_k \cdot S(p_k) \quad (1)$$

但し、 K は予測次数、 a_k は符号化済みの参照画素 p_k ($k = 1, 2, \dots, K$) に関する予測の重み (予測係数) である。 $K = 20$ 次の場合の画素配置の例を図 1 に示す。本方式では符号化済み近傍画素の予測誤差を用いたコンテキストモデリング (後述) によって注目画素を 16 通りのグループのいずれかに分類し、対応するグループの確率モデルに応じた適応的なエントロピー符号化を実行している。ここで、グループ毎に観測された予測誤差信号 e の確率密度関数 $P(e)$ が分散 σ_n^2 ($n = 0, 1, \dots, 15$) のガウス関数でモデル化できるものと仮定する。

$$P(e) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \cdot \exp\left(-\frac{e^2}{2\sigma_n^2}\right) \quad (2)$$

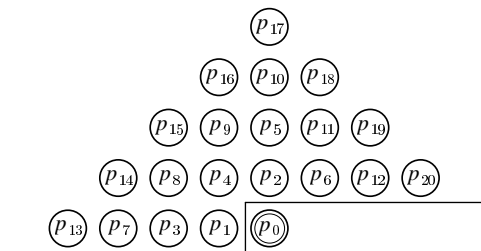


図 1 画素配置 ($K = 20$)

上式より、符号化対象領域 R において予測誤差 e の符号化に必要なビット数を、以下の情報量 $I(R)$ で見積もることができる。

$$I(R) = -\sum_{p_0 \in R} \log_2(P(e) \cdot \Delta e) \\ = -\sum_{p_0 \in R} \log_2\left(\frac{\Delta e}{\sqrt{2\pi\sigma_n^2}}\right) + \frac{\log_2 \epsilon}{2} \sum_{p_0 \in R} \frac{e^2}{\sigma_n^2} \quad (3)$$

但し、 Δe は予測誤差 e に対する量子化ステップ幅、 ϵ は自然対数の底を表している。上式の右辺第 1 項は e に依存しない定数であり、第 2 項は各グループに対応した分散 σ_n^2 の逆数で重み付けされた予測誤差の 2 乗和に比例している。このような重み付き 2 乗誤差の最小化により、領域 R に関する符号化レートをもっとする予測器 (予測係数 a_k の組) が設計できる。実際には、画像を 8×8 画素のブロック単位で M 通りのクラスに分類し、クラス毎の予測器の設計とクラスの更新を交互に繰り返すことで、ブロック単位の適応予測に適した予測器 (M 種類) を求めている。なお、予測器の種類 M と予測次数 K の値は画像サイズに応じて適切に設定している [3]。

3. コンテキストモデリングと予測誤差の符号化

本方式では予測誤差 e の量子化代表値を、 $|e|$ が小さい順に割り当てられた非負の整数値 E (量子化インデックス) で表現し、この E に対してエントロピー符号化を適用している。可逆符号化では量子化ステップ幅が $\Delta e = 1$ となるため、8bit 精度の画像では E の値も $0 \sim 255$ の整数値となる [4]。また、予測誤差の絶対値 $|e|$ とインデックス E の間には $E \approx 2 \cdot |e|$ の関係がほぼ成立する。従って、各グループの e の確率密度関数が (2) 式のガウス関数で表せるとき、インデックス E の生起頻度も以下の片側ガウス分布に従うと考えられる。

$$P_n(E) = \alpha_n \cdot \exp\left(-\frac{E^2}{8\sigma_n^2}\right) \quad (4)$$

但し、 α_n は確率の和を 1 とするための係数である。文献 [1] では、(4) 式と 16 通りの σ_n^2 の値に基づいて、上記のインデックス E に対する可変長符号 V_n ($n = 0, 1, \dots, 15$) を設計している。但し、各 σ_n^2 の値は可変長符号 V_n の理論的な平均符号長が $2.0 \sim 7.5$ bits/sample の間で等間隔となるように前もって設定している。また、符号化済みの近傍画素における E の和をしきい値 $Th_1, Th_2, \dots, Th_{15}$ を用いて 16 レベルに量子化することで各画素のコンテ

*東京理科大学 理工学部 電気電子情報工学科

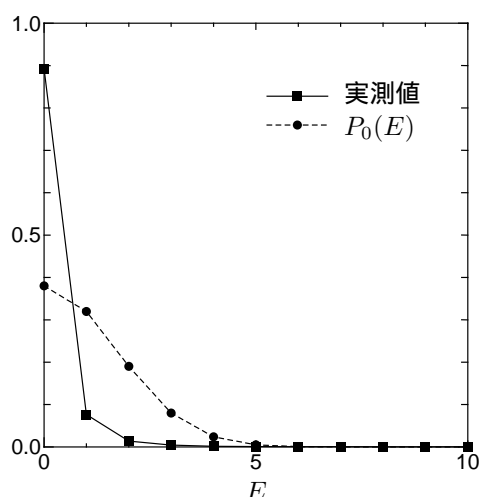


図 2 E の頻度分布の例 (Shapes)

クストモデリング (グループ分類) を実現しており、これにより各グループに対応した可変長符号 V_n を画素単位に適応選択することが可能となる。図 2 は、上記のコンテキストモデリングによって分散最小のグループ ($n=0$) に分類された画素について求めた E の頻度分布の例を示している。ここで使用した Shapes は単純なオブジェクトで構成された CG 画像であり、平坦部で $E=0$ となる頻度が極めて高くなることから、想定した確率モデル $P_0(E)$ との間に大きなミスマッチが生じていることがわかる。

4. 算術符号化の導入

上記の問題を解決するためには実際の頻度分布に応じた適切な σ_n^2 の値を設定する必要があるが、1bit/sample 以下のビット割り当てが原理的に不可能な可変長符号を用いた場合、 σ_n^2 の値が極端に小さいグループでは符号の効率が大きく低下してしまう。そこで、本稿では情報源の拡大を行うことなく柔軟なビット割り当てが可能な算術符号化の導入を図る。本稿で使用するレンジコード [2] は実装が容易であり、また比較的高速な多値算術符号として最近注目を集めている方式である。本方式では、可変長符号の場合と同様なコンテキストモデリングを採用し、グループ毎に (4) 式で与えられる確率モデルとレンジコードを用いて量子化インデックス E の値を符号化している。なお、(4) における σ_n^2 の値は、各グループの理論的なエントロピーが 0.1~7.0 bits/sample の間で等間隔となるように設定した。

5. 特性評価とまとめ

計算機シミュレーションより求めた各方式の符号化レートを表 1 に示す。但し、「Arith.」は算術符号 (レンジコード) をベースとした方式、「Huff.」は可変長符号 (ハフマン符号) を用いる従来の方式 [1] である。また、「TMW」

は文献 [5] の方式、「JPEG-LS」は可逆符号化の国際標準方式 [6] をそれぞれ表している。同表より、「Arith.」は「Huff.」に比較して平均で 0.072 bits/pel 低い符号化レートを達成している。特に、CG 画像である「Shapes」については符号量を 40% 以上削減することができた。これは、予測誤差が 0 に集中する領域において適切なモデリングと符号化効率の改善が実現されたためであると考えられる。また「Arith.」を他の符号化方式と比較すると、「TMW」に比較して 0.009 bits/pel、「JPEG-LS」に比較して 0.308 bits/pel それぞれ低い平均符号化レートを達成していることが確認できる。なお、図 2 からわかるように、画像によっては予測誤差の確率モデルがガウス関数で近似できない場合があり、これが本方式の符号化性能を低下させる一因となっている。今後は、一般化ガウス関数等を用いた適切な確率モデル [7] の導入やコンテキストモデリングの高精度化について検討を加え、更なる性能向上を図る予定である。

表 1 符号化レートの比較 (bits/pel)

Image	Arith.	Huff.	TMW	JPEG-LS
Airplane	3.643	3.661	3.601	3.817
Baboon	5.708	5.727	5.738	6.037
Balloon	2.641	2.653	2.649	2.904
Barb	3.888	3.911	4.084	4.691
Barb2	4.282	4.323	4.378	4.686
Camera	4.071	4.093	4.098	4.314
Couple	3.495	3.525	3.446	3.699
Goldhill	4.256	4.291	4.266	4.477
Lena	4.323	4.348	4.300	4.607
Lennagrey	3.941	3.961	3.908	4.238
Noisesquare	5.386	5.365	5.542	5.683
Peppers	4.251	4.272	4.251	4.513
Shapes	0.994	1.689	0.740	1.214
Average	3.914	3.986	3.923	4.222

【参考文献】

- [1] 松田, 森, 前田, 伊東: 可逆符号化のためのレートを最小とする予測器の設計と評価, 信学論 (D-II), Vol.J85-D-II, NO.3, pp.448-456, Mar. 2002.
- [2] <http://www.compressconsult.com/rangecoder/>
- [3] 白井, 松田, 伊東: 画素/ブロック適応予測を用いた画像の可逆符号化, 信学技報, IE2001-233, Vol.101, No.722, pp.9-16, Mar. 2002.
- [4] 谷口, 上倉, 伊東, 宇都宮: 高階調画像の可変長符号選択形可逆予測符号化方式, 信学論 (B), Vol.J70-B, No.6, pp.654-663, Jun.1987.
- [5] B. Meyer and P. Tischer: TMW a New Method for Lossless Image Compression, Proc. of PCS'97, pp.533-538, Sep. 1997.
- [6] ISO/ITE 14495-1, Information Technology Lossless and Near-lossless Compression of Continuous-Tone Still Images baseline, 1998.
- [7] 松田, 本橋, 伊東: 対象画像毎に予測器と可変長符号を反復最適化する可逆符号化, 第 1 回情報科学技術フォーラム (FIT) 講演予稿集, Sep. 2002.