LJ-12

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

Lossless Coding Using Predictors and VLCs Iteratively Optimized for Each Image

松田　一朗　　本橋 毅　　伊東　晋
Ichiro MATSUDA　　Tsuyoshi MOTOHASHI　　Susumu ITOH
東京理科大学　理工学部
Faculty of Science and Technology,　Tokyo University of Science

1. まえがき

予測符号化に基づいた画像データの可逆圧縮方式では，予測器の効率が符号化性能（圧縮率）を左右する重要なファクタとなる．最近では，対象画像毎に線形予測器の最適化（トレーニング）を行う方式 [1] が，その圧縮率の高さから注目を集めている．通常，予測器の最適化は予測誤差電力 (MSE) をコスト関数として行われるが，符号量の削減を最終目的とする可逆符号化では MSE の最小化が最良の結果を与えるとは限らない．これに対し，筆者らは予測誤差の情報量をガウス分布モデルに基づいて定式化し，その最小化に基づいて可逆符号化に適した予測器を設計する方法を提案した [2]．この方法で求めた予測器は，MSE を最小とする予測器（MMSE 予測器）よりも優れた符号化性能を示すが，予測誤差の頻度分布が上記モデルに従わない場合は必ずしも最適な性能が得られない．本稿では，予測誤差に対する可変長符号を一般化ガウス分布に基づいて設計すると共に，その平均符号長が最小となるよう分布形のパラメータや予測器等を繰り返し最適化する手法について検討する．

2. 予測誤差の情報量と予測器の設計

画素 $\boldsymbol{p}_k$ の輝度値を $S(\boldsymbol{p}_k)$ と表記したとき，注目画素 $\boldsymbol{p}_0$ に関して線形予測を実行した際の予測誤差 e は次式で表せる．

$$e = S(\boldsymbol{p}_0) - \sum_{k=1}^{K} a_k \cdot S(\boldsymbol{p}_k) \tag{1}$$

但し，K は予測次数，a_k は符号化の終了した画素 $\boldsymbol{p}_k$ $(k = 1, 2, \cdots, K)$ の注目画素 $\boldsymbol{p}_0$ に対する予測の重み（予測係数）である．図 1 に $K = 20$ 次の場合の画素配置の例を示す．本方式では予測誤差 e を符号化する際に，符号化済み近傍画素のコンテクストモデリングに基づいて注目画素を 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て可変長符号 V_n $(n = 0, 1, \cdots, 15)$ を適応的に選択している [2]．ここで，グループ毎に観測された予測誤差 e の確率密度関数 $P(e)$ が分散 σ_n^2 $(n = 0, 1, \cdots, 15)$ のガウス関数でモデル化できるものと仮定すると，符号化対象領域 $\boldsymbol{R}$ に関して予測誤差 e の符号化に必要なビット数を，以下の情報量 $I(\boldsymbol{R})$ で見積もることができる．

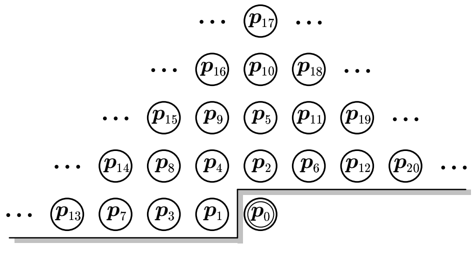

図 1　画素配置

$$\begin{aligned} I(\boldsymbol{R}) &= -\sum_{p_0 \in \boldsymbol{R}} \log_2 \bigl(P(e) \cdot \Delta e\bigr) \\ &= -\sum_{p_0 \in \boldsymbol{R}} \log_2 \left(\frac{\Delta e}{\sqrt{2\pi\sigma_n^2}} \right) + \frac{\log_2 \epsilon}{2} \sum_{p_0 \in \boldsymbol{R}} \frac{e^2}{\sigma_n^2} \end{aligned} \tag{2}$$

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．このような重みつき 2 乗誤差を最小とする予測器（予測係数）は，Wiener-Hopf 方程式に類似した正規方程式を解くことで容易に求めることができ，その結果として領域 $\boldsymbol{R}$ に関する符号化レートを効果的に削減できると考えられる．

実際の画像ではエッジの方向やテクスチャの性質が異なる領域が混在しているため，複数の予測器を用いた適応予測が有効である．本方式では文献 [3] と同様に画像をブロック単位で複数のクラスに分類し，同一のクラスに属するブロックの集合に対して上記の手順を適用することで，適応予測に適した複数の予測器を設計している．

3. 可変長符号の最適化

予測誤差 e の量子化出力を $|e|$ が小さい順に割り当てた非負の整数（インデックス E）で参照するものとすると，原信号のダイナミックレンジ（8 bit 精度の画像の場合 0〜255）の制約を受けない範囲で $E \approx 2 \cdot |e|$ の関係が成立する．従って，各グループにおける予測誤差 e の確率密度関数が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．

$$P_n(E) = \alpha_n \cdot \exp\left(-\frac{E^2}{8\sigma_n^2} \right) \tag{3}$$

但し，α_n は確率の和を 1 とするための係数である．文献 [2] では，予め与えられた 16 通りの σ_n の値と (3) 式に基づいて，上記のインデックス E に対する可変長符号 V_n $(n = 0, 1, \cdots, 15)$ を設計している．また，先に述べた画素毎のコンテクストモデリング（グループ分類）は，符号化済みの近傍画素における E の和をしきい値 $Th_1, Th_2, \cdots, Th_{15}$ を用いて 16 レベルに量子化することにより実現されている．図 2 は，画像 Camera についてグループ毎の E の頻度分布を調べた結果である．これより，上記のコンテクストモデリングによって予測誤差のアクティビティ（分散）はある程度推定できるが，個々のグループにおける E の頻度分布に関しては片側ガウス関数が必ずし

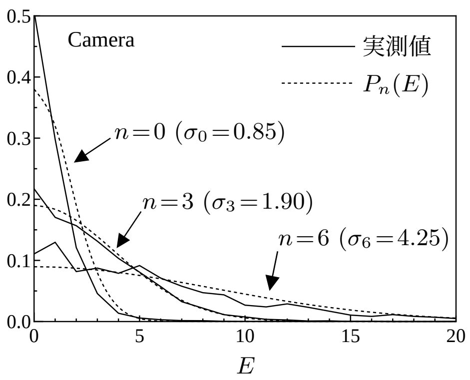

図 2 量子化インデックス E の頻度分布

も適切なモデルとはいえないことがわかる．このような分布形の不一致に起因した符号化効率の低下を防ぐため，本方式では可変長符号を設計する際の確率モデルとして新たに一般化ガウス関数 $G_{c,n}(E)$ を導入している．

$$G_{c,n}(E) = \alpha_n \cdot \exp\left(-\left|\sqrt{\frac{\Gamma(3/c)}{\Gamma(1/c)}} \cdot \frac{E}{2\sigma_n}\right|^c\right) \qquad (4)$$

ここで c は一般化ガウス関数の性質を決めるパラメータであり，上式は $c=1$ のときラプラス関数，$c=2$ のときガウス関数にそれぞれ一致することが知られている．本稿ではパラメータ c のとり得る値を 0.2〜3.2 の間（0.2 刻み）とし，これらの中で可変長符号の効率が最大となる値をグループ毎に一つ選択している．実際には，グループ毎に 16 通りの生起確率 $G_{c,n}(E)$ $(c=0.2, 0.4, \cdots, 3.2)$ に対応した可変長符号 $V_{c,n}$ をそれぞれ求めておき，E の頻度分布を用いて計算した $V_{c,n}$ の平均符号長が最小となるときの c の値を選択する．最終的な c の値は，グループ毎に 4 ビットの付加情報（計 64 ビット）として符号化され，復号側ではこの情報に基づいて各グループで使用される可変長符号を用意する．

4. 予測器の最適化

可変長符号 $V_{c,n}$ の設計に用いる確率モデルとして非ガウス関数 $(c \neq 2)$ が選択された場合，**2.** で述べた予測器の設計に際しても，同様な確率モデルを用いることにより符号化性能を更に改善できるものと予想される．しかし，予測誤差の確率密度関数 $P(e)$ としてガウス関数以外を仮定すると，情報量 $I(\boldsymbol{R})$ を (2) 式のような重み付き 2 乗誤差として扱うことが不可能となり，その最小解を一意に求めることも極めて困難となる．そこで，本稿では可変長符号 $V_{c,n}$ を用いて算出された符号量をコスト関数とし，予測係数 a_k の値を直接修正する方法を提案する．また，この修正処理によってコンテクストモデリングのしきい値や c の値，予測器の選択状況（クラス分類）といったパラメータの最適値も変動するため，これらに関する最適化処理を更に繰り返すことで画面全体の符号量を最小化する．具体的な処理手順を以下に述べる．

(1) **2.** の手順に従って複数の予測器および各種パラメータの初期値を決定する．

(2) 各クラスについて求めた符号量をコスト関数として予測係数を修正する．この修正処理は乱数で選んだ 2 つの予測係数 a_i, a_j を組として繰り返し実行される．また，修正後の予測係数を $a'_i = a_i + \Delta a_i$, $a'_j = a_j + \Delta a_j$ とおくと，予測誤差も $e' = e - \{\Delta a_i \cdot S(\boldsymbol{p}_i) + \Delta a_j \cdot S(\boldsymbol{p}_j)\}$ と表せることから，$\Delta a_i, \Delta a_j$ の探索範囲（予測係数の量子化精度の ±5 倍）に対応する更新項（{ } 内の値）を求めることで，e' の計算に必要な演算量を抑えている．

(3) 各クラスについて符号量を最小化するしきい値 $Th_1, Th_2, \cdots, Th_{15}$ の組合せを動的計画法 [2] を利用して決定する．

(4) 各グループについて可変長符号 $V_{c,n}$ の平均符号長を最小とするパラメータ c を決定する．

(5) 各ブロックについて符号量を最小化する予測器（クラス）を選択する．

(6) (2)〜(5) の手順を画面全体の符号量が減少しなくなるまで繰り返す．

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本方式の符号化レートを表 1 に示す．但し,「情報量」は予測誤差の情報量 $I(\boldsymbol{R})$ の最小化に基づいて予測器を設計する方式 [2],「MMSE」は予測誤差電力（MSE）の最小化に基づいて予測器を設計する方式である．これら 2 方式については予測器と可変長符号が異なることを除いて本方式と同一の符号化手順が適用されている．また,「TMW」は文献 [1] の方式,「JPEG-LS」は可逆符号化の国際標準方式 [4] をそれぞれ表している．同表より，可変長符号と予測器を実際の符号量に基づいて最適化する本方式は,「情報量」に比較して 0.017 bits/pel,「MMSE」に比較して 0.041 bits/pel それぞれ低い符号化レートを達成していることが確認できる．なお，本方式の符号化性能は画像によって「TMW」よりも劣る場合があるが，これは「TMW」が算術符号化をベースにしており，平坦部などで 1 bit/pel 以下の符号化が可能であるのに対し，可変長符号を用いる本方式では原理的に 1 bit/pel 以下の符号化が実現できないためであると考えられる．今後は「TMW」より復号処理が簡易であるという本方式の特徴を保持しつつ，算術符号を導入する予定である．

表 1 符号化レートの比較 (bits/pel)

Image	本方式	情報量	MMSE	TMW	JPEG-LS
Airplane	3.650	3.663	3.688	**3.601**	3.817
Baboon	**5.723**	5.730	5.759	5.738	6.037
Balloon	**2.642**	2.654	2.660	2.649	2.904
Barb	**3.897**	3.916	3.929	4.084	4.691
Barb2	**4.302**	4.320	4.339	4.378	4.686
Camera	**4.055**	4.093	4.151	4.098	4.314
Couple	3.495	3.525	3.565	**3.446**	3.699
Goldhill	4.279	4.290	4.312	**4.266**	4.477
Lena	4.333	4.344	4.360	**4.300**	4.607
Lennagrey	3.958	3.970	3.980	**3.908**	4.238
Noisesquare	**5.320**	5.365	5.367	5.542	5.683
Peppers	4.262	4.272	4.296	**4.251**	4.513
Average	**4.160**	4.178	4.201	4.189	4.472

【参考文献】

[1] B. Meyer et al.: “TMW - a New Method for Lossless Image Compression”, Proc. of PCS’97, pp.533–538, Sep. 1997.

[2] 松田 他: “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”, 電子情報通信学会論文誌, vol.J85-D-II, no.3, pp.448-456, Mar. 2002.

[3] F. Golchin et al.: “Classified Adaptive Prediction and Entropy Coding for Lossless Coding of Images,” Proc. of ICIP’97, vol.III, pp.110–113, Oct. 1997.

[4] ISO/IEC, ISO/IEC 14495-1:1999, “Information Technology – Lossless and Near-lossless Compression of Continuous-Tone Still Images: Baseline”, Dec. 1999.