

画像に対する状態表現を用いたモデル化と無歪み符号化

State-Based Modeling and Lossless Coding of Images

西村 信哉*
Shinya NISHIMURA

堀井 俊佑*
Shunsuke HORII

松嶋 敏泰*
Toshiyasu MATSUSHIMA

Abstract— We consider the modeling and lossless coding of continuous-tone images. We call the modeling state-based, when the distribution of sample value is defined by a conditional probability given a state. We use this modeling for Bayes coding. In a sense, the existing lossless coding algorithms and lossy codings using the transform can be formulated as the form of state-based image modeling. So, we consider some modeling schemes and survey the compression ratio of various modelings experimentally. Then we discuss some ways of modeling, which include to identify a state, to decide a predictor and to set prior distribution, as the indicator of effective lossless image coding.

Keywords— lossless image coding, state-based image modeling, Bayes coding

1 はじめに

本研究では、濃淡画像に対する確率モデル化と無歪み符号化を取り扱う。

一般的に、無歪み符号化の中でも特に、確率モデルに基づく符号化はモデル化とそのモデルの下での符号化確率を計算することで行われる。 t 時点の画素値 x_t に対して、それ以前に観測された x_1^{t-1} の条件付確率 $P(x_t | x_1^{t-1})$ を計算し、符号化ではこの条件付確率を用いてエントロピー符号化を行う。

また無歪み符号化法の研究の一つに Bayes 符号 [1] がある。これは状態によって確率分布に対してその事前分布のクラスで重み付けることによって平均的に最適な符号長を与えるものである。

濃淡画像に対する無歪み符号化は、これまでに LOCO-I[2] や CALIC[3] など様々な方式が提案され、国際標準規格である JPEG-LS など実用的に利用されている。これらは符号化以前の画素値（特に近隣の画素）から符号化したい画素の予測値を計算し、その誤差を周辺の画素値の変化を表現した誤差分布で 2 値符号化するアルゴリズムである。

一方、濃淡画像に対する有歪み符号化では、フーリエ変換や離散コサイン変換、ウェーブレット変換等の周波数変換を用いて、その変換の出力の偏りと画像の周期性を上手く利用した符号化を行っている。特に国際標準規格である JPEG[4] は離散コサイン変換を画像のサブブロックに適時適応し、得られる係数が低周波成分に大き

な値が出ることを用い、比較的劣化の影響が少ない高周波成分の係数をカットすることで、歪みを許して高い圧縮率を達成している。

本研究ではこれらの符号化の特色を考え、濃淡画像に対する無歪み符号化の手法である予測値とその誤差を用いる符号化を基本とし、予測値に周波数変換によって得られる係数値を用いる。予測値は複数の予測式から算出し、それぞれが誤差分布を持つ状態とみなす。その状態クラスに対して、ベイズ符号を行い、このようなモデルクラスの評価を行う。

2 対象とする画像

画像は横 N サンプル、縦 M サンプルのサイズを持つ濃淡画像とし各サンプルの画素値は $\mathcal{X} = \{0, 1, \dots, 255\}$ のいづれかの値をとるものとする。符号化ではラスタスキャンで順に $n (= N \times M)$ サンプルが 1 つずつ符号器に入力される (図 1)。符号化に際して、各サンプルの画素値及び横 N 、縦 M のサイズは既知とする。

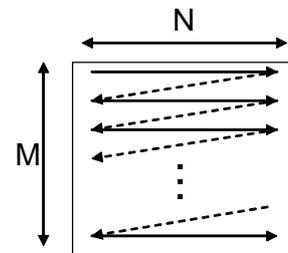


図 1: 対象とする画像と符号化順

3 状態表現を用いた情報源モデル

以下では本研究で中心となるモデル化を示す。予測値と誤差の関係を整理し、予測の仕方が状態となる誤差分布を定義し、その状態表現を用いたモデルに対するベイズ符号の定式化を行う。

3.1 予測値と予測誤差

画素値には次式のような関係があると仮定する。

$$x_t = h(x_1^{t-1}) + \epsilon_t. \quad (1)$$

これは、目的変数 x_t の値を、それ以前の $t-1$ 個の説明変数 x_1^{t-1} の関数と誤差 $\epsilon_t \in \{-255, \dots, 255\}$ で表現し

* 〒 169-8555 東京都新宿区大久保 3-4-1 早稲田大学 Waseda University Ohkubo 3-4-1, Shinjuku-ku, Tokyo, 169-8555 Japan.
E-mail: nishimura@matsu.mgmt.waseda.ac.jp

たものである．すなわち各符号化時点 t において予測値 \hat{x}_t が，

$$\hat{x}_t = h(x_1^{t-1}), \quad (2)$$

として定められる．

このとき，誤差分布に以下の多項分布を仮定する．

$$P_e(i) = \theta_i, \quad -255 \leq i \leq 255. \quad (3)$$

θ_i が i の出現確率である．これを用いて以下の確率モデルが定義される．

$$P(x_t | x_1^{t-1}, \theta) = P(x_t - \hat{x}_t). \quad (4)$$

3.2 予測の仕方を状態とする誤差分布

予測関数 $h_s(x_1^{t-1})$, $\forall s \in \mathcal{S}$ で \hat{x}_t が計算されるとき，誤差分布（パラメータ θ^s ）がそれぞれ定義される．

$$P(x_t | x_1^{t-1}, \theta^s, s) = P(x_t - \hat{x}_t). \quad (5)$$

このインデックス s を各分布を決定する状態とみなし， \mathcal{S} がそのときの状態空間とする．

3.3 状態表現を用いたモデルに対するベイズ符号

状態空間 \mathcal{S} を持つモデルに対して，ベイズ符号の符号化確率は以下の式で与えられる．

$$P_c(x_t | x_1^{t-1}) = \sum_{s \in \mathcal{S}} \int P(x_t | x_1^{t-1}, \theta^s, s) \times P(\theta^s | x_1^{t-1}, s) d\theta^s P(s | x_1^{t-1}). \quad (6)$$

ただし，各状態とパラメータの事前分布 $P(s)$, $P(\theta^s | s)$ は既知であるとする．

これは，複数の状態をその事前分布とパラメータの事前分布で重み付けたモデルになっており，平均的に冗長度を最小にする最適な符号長を与える符号化確率である．

ここで，誤差分布に多項分布を考えているため，パラメータの事前分布 $P(\theta^s | s)$ にディリクレ分布（パラメータ $\alpha(-255), \dots, \alpha(255)$ ）を仮定することで，式 (6) は以下のように計算できる．

$$P_c(x_t | x_1^{t-1}) = \sum_{s \in \mathcal{S}} \frac{\eta(\epsilon_t | s) + \alpha(\epsilon_t)}{\sum_{i=-255}^{255} (\eta(i | s) + \alpha(i))} P(s | x_1^{t-1}). \quad (7)$$

ただし， $\eta(i | s)$ は状態 s の下での i の出現回数である．すなわちディリクレ分布のパラメータ $\alpha(i)$ はカウンタの初期値と等価である．

4 周波数変換を用いた予測

JPEG 等に用いられている周波数変換は，画素値を定められた基底の係数値へ変換することで与えられる．こ

れは基底の選び方によって，画像の領域ごとに周期的に画素値が変化するようなモデルとみなすことができ，その係数がそれぞれの周期の大きさを表す量になっている．JPEG では離散コサイン変換を，JPEG2000 ではウェーブレット変換をそれぞれ利用している．

ここでは基底とその係数で決まる予測の対応と，JPEG に用いられる離散コサイン変換の本研究における着目点を示す．

4.1 変換系とその確率モデル化

任意の変換 T は基底 $f_j (\forall j \in \mathcal{J})$ を持ち，長さ $|\mathcal{J}|$ の基底ベクトルを $(f_1, \dots, f_{|\mathcal{J}|})$ ，その係数ベクトルを $(a_1, \dots, a_{|\mathcal{J}|})$ としたとき実現値（画素値） x_t は，

$$x_t = \sum_{j \in \mathcal{J}} a_j f_j(t) = (a_1, \dots, a_{|\mathcal{J}|}) \cdot (f_1, \dots, f_{|\mathcal{J}|}), \quad (8)$$

で表されるベクトルの内積で定義される．ただし， \cdot は内積を表す記号とする．

今，基底ベクトル $(f_1, \dots, f_{|\mathcal{J}|})$ を固定した場合，式 (8) より x_t は t 時点までに得られる変換係数ベクトル $(a_1, \dots, a_{|\mathcal{J}|})_t$ によって一意に定まる．

しかし，符号化を考えると， t 時点においては x_t の値が未知なので，それ以前の画素値 x_1^{t-1} から変換 T によって求められる $(a_1, \dots, a_{|\mathcal{J}|})_{t-1} = T(x_1^{t-1})$ を用いて $\hat{x}_t = (a_1, \dots, a_{|\mathcal{J}|})_{t-1} \cdot (f_1, \dots, f_{|\mathcal{J}|})$ などのように予測することになる．ここで係数から予測値を求める関数 g によって，

$$\begin{aligned} \hat{x}_t &= h(x_1^{t-1}) \\ &= g(T(x_1^{t-1})) \\ &= g((a_1, \dots, a_{|\mathcal{J}|})_{t-1}), \end{aligned} \quad (9)$$

と表現でき，この予測誤差に多項分布を仮定する．

4.2 離散コサイン変換

離散コサイン変換（以下，DCT）は，変換の基底として正弦波を低周波なものから高周波なものまで変換のサイズ分利用する．そして得られた係数列はそれぞれの周波数成分の大きさ（振幅）を表す．

JPEG では，これを 2 次元に拡張した 2 次元 DCT を用いて符号化を行う．図 2 左が， 8×8 の 2 次元 DCT の基底画像である．周波数成分が横，縦方向に 2 通りあり，左から右へ，及び上から下へ周波数が高くなっている．図 2 右は，変換によって得られる係数行列を示しており，係数の位置と基底のパターンの位置はそれぞれ対応している．係数 $a_{0,0}$ は画素の平均，係数 $a_{0,1}$ は横方向のグラデーション，係数 $a_{1,0}$ は縦方向のグラデーション，係

数 $a_{1,1}$ は斜め方向のグラデーション, ..., 係数 $a_{7,7}$ は格子状のパターンの大きさをそれぞれ表している.

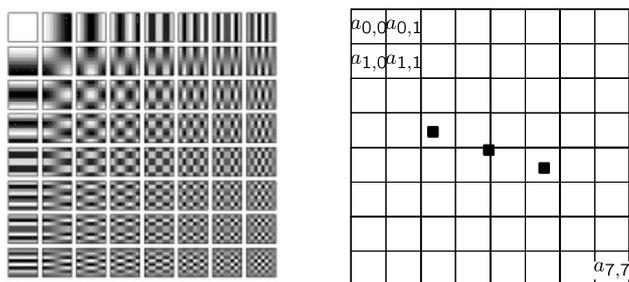


図 2: 2 次元 DCT の基底と係数行列

4.3 DCT と符号化の関係

JPEG では, 上述の係数行列において画像の周期性を表現すると共に, 低周波 (左上) の係数が大きな値をとりやすく高周波 (右下) の係数が小さい値をとりやすいという画像データの持つ性質に着目し, 画像からサブブロックを切り出し, 量子化と係数行列の値の小さい高周波の係数値を切り捨てとによって, 圧縮符号化を達成する. しかし, これでは歪みを発生させ元の画像に復元することができない.

そこで本研究では, DCT 係数行列において全体的な特徴を表す低周波成分に大きな値が出現する性質を予測値の計算に利用する. そして複数の予測式が状態となるモデルに対してベイズ符号を用いて 2 値符号化することで無歪み符号化を行う.

例えば, $|a_{1,0}| = \max(|a_{0,1}|, |a_{1,0}|, |a_{1,1}|)$ であるとき, 基底 $f_{1,0}$ のパターンである縦方向のグラデーションが強いということを示しており, この状態においては左右の画素値間が上下の画素値間より近い値になると言える. この状態では, 予測値を左隣の画素を用いて決定することが考えられる.

$$\hat{x}_t = x_{t-1}, \text{ if } |a_{1,0}| = \max(|a_{0,1}|, |a_{1,0}|, |a_{1,1}|). \quad (10)$$

5 本研究における符号化アルゴリズム

符号化したい画像の各画素値と横・縦のサイズは既知とする. また状態・誤差分布の事前分布も既知である.

5.1 ベイズ符号化アルゴリズム

1. 画像端の画素から DCT 係数を計算
2. t : 符号化する時点の画素 x_t を読み込む
3. 得られた DCT 係数からそれぞれの状態 s を定め, そのときの予測値 \hat{x}_t を計算する
4. $\epsilon_t = x_t - \hat{x}_t$ を状態の分布と誤差分布に従ってベイズ符号化
5. DCT 係数の更新 - x_t を用いて新たに DCT 係数を計算しなおす

6. 誤差分布の更新 - カウンタを用いて学習

7. 状態の事後確率を更新

8. $t < n$ のとき 2 に戻る.

5.2 DCT のサイズ

DCT 基底の一辺の大きさ K を $K = 4, 8, 16$ の 3 種類調べる. これらの違いは近い領域の関係を見るか遠い領域の関係を見るかの違いとなる.

5.3 誤差分布の事前分布

誤差分布を考えているので, $\epsilon = 0$ 近辺の確率が高い分布を用いる. 事前分布の設定で前述のような性質を表現する. そのときのディリクレ分布のパラメータ (カウンタの初期値) のパターンが複数考えられる. 図 3 では, 2 種類の事前分布を重ねて表示した.

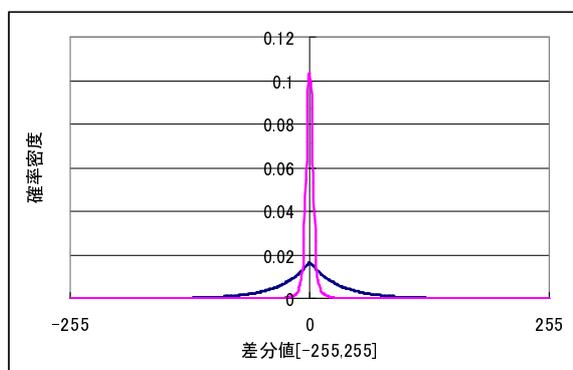


図 3: 誤差の事前分布

5.4 予測式

予測値関数の決め方の種類は無数あるが, その効用が説明しやすい係数行列と期待される画素値の対応が明確なもので且つ計算量が少ないものが望まれる.

ここでは DCT 係数行列の内, 左上の 4 つの係数値組 $(a_{0,0}, a_{0,1}, a_{1,0}, a_{1,1})_{t-1}$ を周期を表す指標として状態に変換し利用する. この 4 つの係数値の値は特に大きく, これらのパターンが画素値の基本的な性質を与える. 他の係数値に表れる小さな値は誤差として分布に表現する. またこの 4 つの係数値だけを計算するアルゴリズムにより計算量も大きく削減できる.

以下, $A = \max(|a_{0,1}|, |a_{1,0}|, |a_{1,1}|)$ とし, 様々な性質を持つ予測式を挙げる.

5.4.1 予測式 1

全体の平均が基本になっていて, その平均とのずれが誤差になる 1 次回帰モデルである.

$$\hat{x}_t = \left\lfloor \frac{a_{0,0}}{K} \right\rfloor. \quad (11)$$

5.4.2 予測式 2

グラデーションの方向（周波数成分のパターン）を係数の大小で決定し、その向きで近い値になるであろう近隣の値を用いる。周囲の傾向から状態を逐次切り替えて用いる（従来の予測符号化と同等）。

$$\hat{x}_t = \begin{cases} x_{t-N} & , A = |a_{0,1}| \\ x_{t-1} & , A = |a_{1,0}| \\ x_{t-N-1} & , A = |a_{1,1}|. \end{cases} \quad (12)$$

5.4.3 予測式 3

予測式 2 の向きで、もう一つ隣の画素との平均を用いる。

$$\hat{x}_t = \begin{cases} \left[\frac{x_{t-N} + x_{t-2N}}{2} \right] & , A = |a_{0,1}| \\ \left[\frac{x_{t-1} + x_{t-2}}{2} \right] & , A = |a_{1,0}| \\ \left[\frac{x_{t-N-1} + x_{t-2N-2}}{2} \right] & , A = |a_{1,1}|. \end{cases} \quad (13)$$

5.4.4 予測式 4

それぞれの係数の絶対値が、周期の特性を表す重みになっていると見なして、それぞれの予測値（予測式 2）を重み付けて用いる。

$$\hat{x}_t = \left[\frac{|a_{0,1}| \times x_{t-N} + |a_{1,0}| \times x_{t-1} + |a_{1,1}| \times x_{t-N-1}}{|a_{0,1}| + |a_{1,0}| + |a_{1,1}|} \right]. \quad (14)$$

5.4.5 予測式 5

それぞれの係数の 2 乗が、周期の特性を表す重みになっていると見なして、それぞれの予測値（予測式 2）を重み付けて用いる。

$$\hat{x}_t = \left[\frac{(a_{0,1})^2 \times x_{t-N} + (a_{1,0})^2 \times x_{t-1} + (a_{1,1})^2 \times x_{t-N-1}}{(a_{0,1})^2 + (a_{1,0})^2 + (a_{1,1})^2} \right]. \quad (15)$$

6 実験

本研究では、[5]、[6] にある 38 枚のテスト画像を用いて符号化を行い、圧縮率 [bits/symbol] で評価を行う。以下の表ではその内、自然画（cafe [2048 × 2560], bird [256 × 256]（図 4 左上）, frog [621 × 498]）、イラスト（circles [256 × 256]（図 4 右上）, text [256 × 256], france [672 × 496]）、それらの混合画像（cmpnd2 [5120 × 6624], us [512 × 448], library [464 × 352]（図 4 下））の計 9 枚の画像に対する結果を示す。

表 1 から 3 は、DCT 基底の一辺のサイズ $K = 4, 8, 16$ ごとに 6 章の予測式をそれぞれ単独で使用した場合とそれらを状態として事前分布（一様）で重み付けたベイズ符号の圧縮率である。表 4 は JPEG で最も劣化の少なくなるように符号化した結果である。

表 1: $K = 4$ のときの、各予測式とベイズ符号化

	予測式 1	予測式 2	予測式 3	予測式 4	予測式 5	ベイズ
café	7.86	5.97	6.25	5.97	5.94	5.97
bird	7.16	4.23	4.49	3.97	4.03	3.97
frog	6.73	5.20	6.12	6.20	6.14	5.20
circles	5.85	0.17	0.30	1.75	1.76	0.17
text	3.53	0.81	1.19	3.40	3.36	0.81
france	7.59	1.01	1.65	2.51	2.37	1.01
cmpnd	7.66	2.07	3.21	2.98	2.84	2.07
us	6.23	3.23	3.68	3.72	3.63	3.23
library	7.78	5.83	6.17	6.20	6.18	5.83

表 2: $K = 8$ のときの、各予測式とベイズ符号化

	予測式 1	予測式 2	予測式 3	予測式 4	予測式 5	ベイズ
café	7.86	5.94	6.24	5.98	5.94	5.94
bird	7.18	4.25	4.52	3.97	4.02	3.97
frog	6.71	5.26	6.15	6.21	6.20	5.26
circles	5.50	0.18	0.31	0.98	1.00	0.18
text	3.54	0.77	1.13	3.14	3.25	0.77
france	7.43	1.09	1.78	2.49	2.29	1.09
cmpnd	7.66	2.02	3.15	2.81	2.62	2.02
us	6.20	3.73	4.01	3.90	3.90	3.73
library	7.72	5.90	6.24	6.31	6.22	5.90

表 3: $K = 16$ のときの、各予測式とベイズ符号化

	予測式 1	予測式 2	予測式 3	予測式 4	予測式 5	ベイズ
café	7.86	5.93	6.17	5.91	5.89	5.91
bird	7.21	4.32	4.59	4.07	4.12	4.07
frog	6.71	5.30	6.15	6.25	6.19	5.30
circles	4.89	0.19	0.34	0.87	0.90	0.19
text	3.22	0.83	1.20	3.04	2.87	0.83
france	7.17	0.96	1.64	2.54	2.39	0.96
cmpnd	7.67	2.03	3.19	2.91	2.74	2.03
us	6.15	3.31	3.76	3.84	3.73	3.31
library	7.56	5.96	6.30	6.32	6.26	5.96

表 4: JPEG による符号化（劣化最小）

café	5.60	circles	3.13	cmpnd2	3.10
bird	5.38	text	10.63	us	4.07
frog	7.08	france	3.98	library	7.20

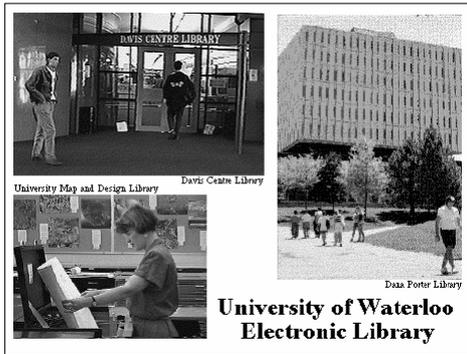


図 4: テストイメージ (bird,circles,library)

7 考察

7.1 DCTのサイズによる違い

これは予測方法と画像の組み合わせによる差があり、一概に優劣はつけ難い結果となった。ただし、例えば $K = 4$ で上手くいくものは近い領域で周期の関係性が高いということができ、その逆もまた言える。

7.2 画像のタイプと予測式の関係

自然画、イラスト、混合の3つのタイプの画像に対して比較すると、自然画には予測式5が、イラストには予測式2と3が、混合では予測式2と5が良い結果を示した。自然画では徐々に画素値が変化する特徴がある。予測式5の係数の2乗の重み付けがちょうどそのような徐々に変化する様子を示していると考えられる。イラストでは画素値が急激に変化する位置(輪郭部)とほとんど変化しない大きな領域がある。そのため予測式2や3のように単純に隣の画素をその傾向に合わせてそのまま用いるのが適合している。混合画像ではそれらどちらの性質も有しているため、予測2と5でよい結果となったと考えられる。

7.3 ベイズ符号の結果

予測式を重み付けて計算したベイズ符号では、それぞれの画像に対して良い予測式を用いた場合の圧縮率に近い値を達成した。符号化後の事後確率を調べると、1つの状態だけがほぼ1に、他がほぼ0に近い値をとるような分布になっている。これは符号化過程においてよく圧縮できる予測式の重みが大きくなるということであり、複数の予測式を組み合わせたときにその優劣を自動的に

学習することになる。

7.4 JPEGとの比較

表4のJPEGはもっとも画質を損なわないモードで符号化したものである。これと比較すると、実験結果は無歪み符号化であるにも関わらず良い結果を示した。これはJPEGに比べて画像からの学習量が大きいためであると思われる。

8 まとめ

本稿では、濃淡画像に対する無歪み符号化のための状態表現を用いた情報源モデルの符号化と考察を行った。特に画像の特徴と周波数変換(2次元DCT)の関係を学習し無歪み符号化することで対応関係が整理でき、周波数変換を状態に利用したモデル化と無歪み符号化への指針を示した。

謝辞

本研究を行うにあたり、数多くの御助言、御支援を賜りました松嶋研究室の各氏に感謝致します。なお、本研究の一部は日本学術振興会科学研究費基盤(C)一般(No.18560391)の援助による。

参考文献

- [1] T.Matsushima, H.Inazumi and S.Hirasawa, "A Class of Distortionless Codes Designed by Bayes Decision Theory," IEEE, Trans., Inf. Theory, vol. 35, pp. 1288-1293, Sep. 1991.
- [2] M. J.Weinberger, G. Seroussi, and G. Sapiro, "The LOCO-I Lossless Image Compression Algorithm: Principles and Standardization into JPEG-LS," IEEE Trans. Image Process., vol. 9, pp. 1309-1324, Aug. 2000.
- [3] X. Wu and N. D. Memon, "Context-Based, Adaptive, Lossless Image Coding," IEEE Trans. Commun., vol. 45, pp. 437-444, Apr. 1997.
- [4] ISO/IEC 10918-1, "Digital Compression and Coding of Continuous-tone Still Images: Requirements and Guidelines", Feb. 1994.
- [5] ftp://ftp.csd.uwo.ca/pub/from_wu/images/
- [6] <http://links.uwaterloo.ca/bragzone.base.html>