

ニューラルネットによって構成された予測器による、 適応的な画像圧縮*

水田 智史

弘前大学理工学部†

1 はじめに

画像圧縮の手法として、予測器を用いて入力画素の近傍の画素値を予測し、原画像の画素値との差をデータとして保存するという手法が広く用いられている。本稿では、ニューラルネットによって個々の画像毎に最適な予測器を構成する新たな手法を提案する。すなわち、個々の画像のもつ特徴は画像によってそれぞれ異なるという点に着目し、個々の画像に応じた予測器を構成することで圧縮率の向上を図る。本手法では、画像毎にニューラルネットの学習が必要なため圧縮処理には通常の手法に比べて長時間を必要とするが、画像の伸長に必要な処理時間は通常の手法と同程度であるため、圧縮の処理時間があまり問題にならない画像ライブラリの構築等の場合に有効であると考えられる。

2 予測器の構成

2.1 予測器の設計

予測画素を1画素とし、画素間の一階微分の情報を考慮に入れるため予測画素近傍の縦、横、ななめそれぞれ2画素を入力画素とする(図1)。

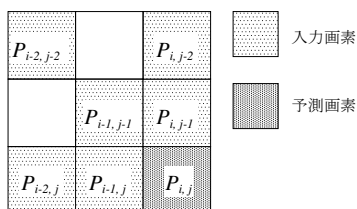


図1: 入力画素と予測画素

ここで、画像(またはブロック)のサイズを $N \times M$ 、座標 (i, j) の原画像の画素値と予測画素値をそれぞれ

$P_{i,j}$ 、 $Q_{i,j}$ 、予測器に対応する関数を g とおくと

$$Q_{i,j} = \begin{cases} g(P_{i-2,j}, \dots, P_{i,j-1}) & \text{for } 2 \leq i \leq N-1, 2 \leq j \leq M-1 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

と表せる。予測が有効にはたれば、 $P_{i,j}$ と $Q_{i,j}$ との差 $D_{i,j} \equiv P_{i,j} - Q_{i,j}$ の分散は $P_{i,j}$ の分散に比べて小さくなるので、適当な符号化を施すことにより可逆圧縮を実現することができる。ここでベクトル量子化等を組み合わせれば、非可逆圧縮を行うことも可能である。また、画像の伸長は $D_{i,j}$ を復号した後、

$$P_{i,j} = \begin{cases} D_{i,j} + g(P_{i-2,j}, \dots, P_{i,j-1}) & \text{for } 2 \leq i \leq N-1, 2 \leq j \leq M-1 \\ D_{i,j} & \text{otherwise} \end{cases} \quad (2)$$

により行う。

2.2 ニューラルネットの構造

予測器として用いるニューラルネットは隠れ層が一層の層状の順方向性ニューラルネットである(図2)。ただし、データが回帰しないという条件下で隠れニューロン間の結合を許す。

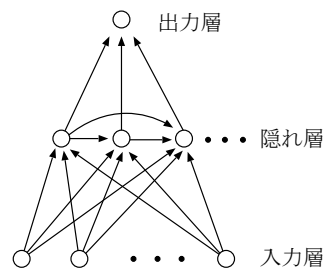


図2: ニューラルネットの構造

* Adaptive Image Compression using Neuralnet Predictor

† Satoshi Mizuta; Faculty of Science and Technology, Hirosaki University

2.3 ニューラルネットの最適化

図 2 に示すニューラルネット全体の空間 $\mathcal{N}$ は、構造空間 $\mathcal{S}$ とパラメータ空間 $\mathcal{P}$ の直積

$$\mathcal{N} = \mathcal{S} \times \mathcal{P} \quad (3)$$

として捉えることができる。ここで、入力層と出力層に属するニューロンの数は予測器のアーキテクチャによって決まっているので、 $\mathcal{S}$ の変数は隠れニューロンの数と結合線で結ぶニューロンのペアであり、 $\mathcal{P}$ の変数は結合線の強度とニューロンの閾値となる。

$\mathcal{S}$ 空間においては

- 隠れニューロンの付加
- 隠れニューロンの削除
- 結合線の付加
- 結合線の削除

を素過程とし、

$$f = N_c + \frac{1}{N_h + R_w}$$

を適応度とした疑似シミュレーテッドアニーリングによる最適化を行う。ここで、 N_c は $D_{i,j} = 0$ を満たす画素数で、 $\mathcal{P}$ 空間における最適化の結果得られる。また、 N_h は隠れニューロンの数、 R_w は全結合の場合の結合線の数に対する実際の結合線の数比である。

$\mathcal{P}$ 空間における最適化はニューラルネットの通常の学習過程に相当し、コスト関数

$$E(\vec{t}) = \frac{1}{2} \sum_{i=2}^N \sum_{j=2}^M |D_{i,j}|^2 \quad (4)$$

が最小になるように $\vec{t}$ を決定する。ここで $\vec{t}$ はすべての結合の強度とニューロンの閾値をベクトルとして表したもので、本研究では共役勾配法を用いた最適化を行った。

なお、この手法は *pruning* や *constructive* な手法 [1] の中間に位置するものである [2]。

3 実験結果

本稿では可逆圧縮について行った実験の結果を示す。

実験は SIDBA の Girl、Lenna、Mandrill の 3 画像をグレースケール (PGM) に変換したものをを用いて行った。原画像は 256×256 のバイナリフォーマットで、ファイルサイズは 65,551 バイトである。 $D_{i,j}$ の符号化には bzip2 を用い、比較の対象として PNG フォーマットを選んだ。実験の結果を表 1 に示す。

表 1 の左端の列の数字はブロックサイズ、bz2 は原画像を直接 bzip2 によって圧縮した結果を示す。size はファイルサイズ (バイト)、ratio は原画像とのファイルサイズの比である。なお、圧縮後のデータには予測器であるニューラルネットのデータを含める必要があるので、ここではそれを 500 バイトと仮定し、単純な圧縮データサイズにそれを加えたものをファイルサイズとしている。

	Girl		Lenna		Mandrill	
	size	ratio	size	ratio	size	ratio
PNG	37198	0.567	39296	0.599	51109	0.780
bz2	42288	0.645	46358	0.707	54922	0.838
8	43198	0.659	46333	0.707	56713	0.865
16	41543	0.634	44330	0.676	55291	0.843
32	40735	0.621	43597	0.665	54782	0.836
64	39895	0.609	43331	0.661	54547	0.832
256	39127	0.597	42236	0.644	54486	0.831

表 1: 実験結果

4 まとめと考察

ニューラルネットを用いた新たな画像圧縮の手法を提案した。SIDBA の標準画像を用いた実験の結果、圧縮比の点に関して、大きなブロックサイズにおいてある程度の効果は得られたものの、現時点において PNG フォーマットによる圧縮には及んでいない。しかし、予測器の精度を上げることでより良い結果が得られる可能性があるものと思われる。

現在、予測器の精度を上げるために、異なった予測器のアーキテクチャやブロック毎に異なる予測器を用いる方法等について検討を行っている。また、今後、ベクトル量子化等と組み合わせた非可逆圧縮についても検討していく。

参考文献

- [1] Tin-Yau Kwok and Dit-Yan Yeung, “Constructive Algorithms for Structure Learning in Feed-forward Neural Networks for Regression Problems”, *IEEE Trans. Neural Networks*, **8** (1997) 630-645; and references therein.
- [2] S. Mizuta, T. Sato, D. Lao, M. Ikeda and T. Shimizu, “Structure Design of Neural Networks Using Genetic Algorithms”, *Complex Systems*, **13** (2001) 161-175.