

ニューラルネットワークによるカラー画像情報圧縮

森田義則, 南 朋之, 袋谷賢吉, 三日市政司, 米田政明

富山大学工学部電子情報工学科計算機工学講座

あらまし Cotrell, Munro, Zipser は, 白黒濃淡画像をいわゆる砂時計型をした3層パーセプトロン状ニューラルネットワークに入力すると, その中間層では, 入力画像の圧縮表現が得られることを示した. その方法を参考にして, 本研究は, 自然カラー画像の情報圧縮を行うことを最終目標としている. しかし, 本方式は Cotrell らと異なり, カラー画像全体を多数の局所画像に分割し, それぞれの局所画像が対応したニューラルネットワークに入力され情報圧縮が行われる形の, 並列分散処理的な構成を持っている. 本稿ではこの構成の内, 局所画像に対する情報圧縮部のみを扱う. 問題を簡単化するため, 入力局所画像は一次元空間画像で, ニューラルネットワークは線形な写像型とした. 実際のカラー画像の例について, このニューラルネットワークの性能および中間層表現を調べた. また, ここで考えているカラー画像の情報圧縮の要請および構成が, 網膜における状況およびその神経構成に類似している点に着目し, 網膜内神経細胞の応答様式と得られたニューラルネットワークの中間層表現を比較した.

Color Image Compression by Neural Networks

Yoshinori MORITA, Tomoyuki MINAMI, Kenkichi FUKUROTANI,

Masashi Mikkaichi, and Masaaki YONEDA

Faculty of Engineering, Toyama University

3190 Gofuku, Toyama-shi, Japan 930

Abstract We proposed here a color image compression method by neural networks. A color image was subdivided into many local images. Each local image was entered into the corresponding neural network. So, the total color image was pararely processed by many local neural networks. The local neural network was three layered perceptron with a few units in the hidden layer which act as image compressors. In this paper, we only analysed and discussed the local neural network which was linear for simplicity. We demonstrated the compressed representaion in the hidden units of the local neural network for one dimensional color image. Further, we noticed that circumstances and architecture of the neural networks are similar to those of retina. We compared the color and spatial weights of the hidden units of the artificial neural networks with spectral and spatial properties of inter-retinal neurons. Many similarities were found between them.

1. まえがき

自然カラー画像は膨大な情報量を含み、その情報圧縮は古くから研究されてきたことであるが、最近の高品位画像に対する社会的要求の高まりから緊急の課題としてクローズアップされてきた。

一般に、与えられたデータを簡単な構造のモデルで近似表現する際に、モデルの構造の決定と近似度の評価が基本的問題となる。エンドユーザーが人間の場合、人間の生理機構や心理物理学の特性がそれらを決める際の参考になる。本稿で扱うカラー画像の情報圧縮の問題に関し、モデルの構造の決定には人間の視覚生理機構が一つの参考になろうし、近似度の評価には人間の視覚心理特性が利用できると思われる。

網膜で捉えられた画像情報は、網膜から伸びた視神経繊維束により、外側膝状体と呼ばれる中継接続所を経て、大脳視覚領へ伝えられる。網膜や大脳視覚領には膨大な数の神経細胞があるのに、視神経繊維の数は人間の場合僅か80万本と言われている。従って、視神経繊維は視覚情報伝達のボトルネックとなっており、このことが、網膜における情報の大幅圧縮の要因となっていると考えられる。

Cottrell, Munro, Zipser は、上に述べた視覚神経系の構成と基本的に似た人工ニューラルネットワークにより、白黒濃淡画像の情報圧縮を行った⁽¹⁾。そのニューラルネットワークは3層のパーセプトロンであり、入力層素子数は64で8x8画素の白黒濃淡画像を入力する。出力層も同数の素子から成り、入力画像を再生する。訓練時の出力教師信号は入力と同じ白黒濃淡画像とする。Cottrell, Munro, Zipser は、中間層の素子数を16にまで減らしても、出力層再生画像の画質が落ちないことを示した。このことは、中間層で入力画像情報の大幅な圧縮が行われたことを意味している。

本研究の目的は2つある。一つはCottrellらの手法のカラー画像への拡張である。もう一つは、そこでの結果からリアルニューラルネットワークすなわち網膜でのカラー画像情報圧縮機構についての示唆を得ることである。

人間の場合、色の恒常性や視覚受容野の変化に見られるように、色彩および空間の表現要素は入力条件により適応的に変化する。人工ニューラルネットワークでそのような機能が実現できるとすれば、それは大きな魅力である。

視覚生理機構にはまだ良く分かっていない点もある。既知の入出力データで写像ニューラルネットワークを訓練し、その中間層の内部表現から、リアルニューラルネットワークの内部機構についての示唆が得られないだろうか。本研究はこれらの目標の第一歩をめざしている。

2. モデル

ここで考えている基本構成は入力画像の局所並列的展開による情報圧縮である。画像再生に際しては、局所情報の他に大域的情報も必要となろうが、本稿ではまず第一歩として、入力画像の局所展開に限って話を進める。

局所情報の圧縮に用いたニューラルネットワークは、入力層、中間の一層および出力層からなる3層のパーセプトロンで、Cottrellらと同じく入力層および出力層の素子数に比べ中間層の素子数の少ない、いわゆる砂時計型である。

入力層の素子と中間層の素子間の結合においては、距離の近いもの同士は結合し易く、遠いもの同士は結合しにくいとする。即ち、結合重みは距離に依存する。ここではユークリッド距離を用い、空間加重としては正規分布を用いる。ここで正規分布を用いたのは、網膜の一次ニューロンである視細胞と、2次ニューロンである水平細胞および双極細胞との結合空間加重からの類推、および、正規分布の標準偏差を小さくした極限がディラックのデルタ関数になるという好ましい性質による。

訓練においては、入力層に画像を与え、出力層には入力画像に正規分布をかけたものを与えた。学習にはバックプロパゲーション法を用い、出力の誤差評価には平均2乗誤差を用いた。

まず、中間層の素子数を1として訓練を行った。次に、その結合重みを固定したまま、中間層の素子数を2とし、この新たに追加した素子のみ同じ入出力データを用いて訓練を行った。さらに、中間層に素子を一つ加え、先の2素子の重みを固定したまま第3番目の素子のみ同じ入出力データを用いて訓練を行った。このようにしたのは、本稿では詳しく述べないが、主成分分析法との関連からである。また、単純化のために中間層および出力層の素子は線形とした。

従って、各素子はアフィン変換を行ったことになる。

3. コンピュータ実験例

入力画像は元来2次元であるが、ここでは簡単化のため一次元画像を用いた。即ち、東大生研の標準画像 girl の 256 x 256 画素から、水平方向 17 画素を取り出した 192 個の一次元画像である。一方、出力層への教師画像は入力一次元画像に正規分布をかけたものである。この入力データを x 、出力データを y と表せば、

$$x = (x_0, x_1, x_2, \dots, x_{16})$$

$$y = (y_0, y_1, y_2, \dots, y_{16})$$

ここで、

$y_i = x_i * \text{EXP}[-(i-8)^2/4]$ $i=0, 1, 2, \dots, 16$ である。

入力画像の R, G, B 各成分についてそれぞれ単独に、上に述べたように一次元入力データを作り、ニューラルネットワークを訓練した。

図 1 に、R, G, B 各成分について、入力層各素子と中間層の第 1 素子、第 2 素子および第 3 素子との結合重みを示した。第 3 素子までで平均 2 乗誤差はいずれも 0.0005 以下となり、かなり良い近似を示している。

ここで、中間層の第 1 素子の結合重みは、R, G, B 各成分共、負の正規分布型となっている。第 2 素子の結合重みは中心が零の左右非対称な奇関数型で、R 成分の場合と G, B 成分の場合では極性が逆になっている。第 3 素子の結合重みは中心部と周辺部で極性の異なる空間対比型となり、やはり、R 成分の場合と G, B 成分の場合では極性が逆になっている。即ち、色対比をも示す 2 重対比型となっている。ここで注目すべき点は、R, G, B 各色彩成分について、第 1 素子の空間重み、第 2 素子の空間重み及び第 3 素子の空間重みの相対形が極性を除きほぼ等しい点である。これより、色彩応答と空間分布とは統計的に独立とみなせる。従って、ニューラルネットワークの色彩部分と空間部分を分けて考える。

次に色彩情報の圧縮を考える。ニューラルネットワークの構成は以下の点を除き上のものと同じである。即ち、入力層と出力層の素子数は共に 3 で、訓練における入力データと出力データは同じであり各画素の R, G, B 成分を与

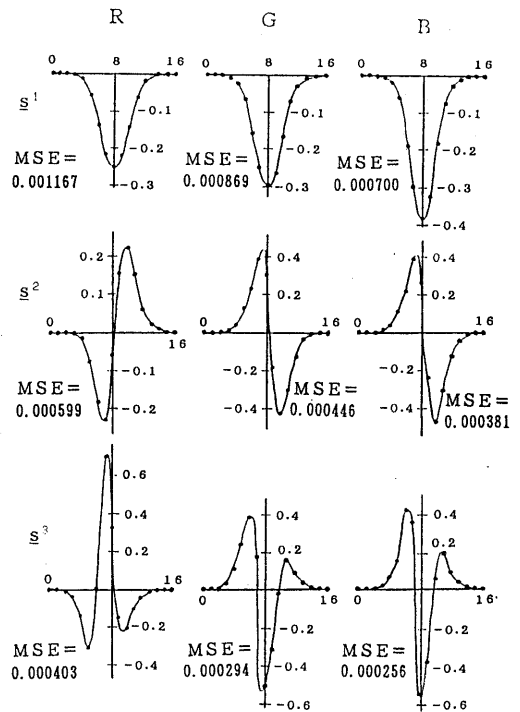


図 1. 一次元画像入力に対する中間層素子の結合重み。R, G, B 列はそれぞれ R, G, B 成分に対するもの。第 1, 第 2, 第 3 行はそれぞれ中間層の第 1, 第 2, 第 3 素子の空間重みを表す。

表 1. 色彩入力に対する中間層素子の結合重みと平均 2 乗誤差

	R	G	B	MSE
第 1 中間素子 (c^1)	0.368 (c_r^1)	0.405 (c_g^1)	0.335 (c_b^1)	0.802E-3
第 2 中間素子 (c^2)	-0.596 (c_r^2)	0.299 (c_g^2)	0.393 (c_b^2)	0.153E-3
第 3 中間素子 (c^3)	0.150 (c_r^3)	-0.898 (c_g^3)	0.858 (c_b^3)	0.1E-5

える。いわゆる恒等写像型である。

表1に、入力層各素子と中間層の第一素子、第2素子および第3素子との結合重み及び素子数を累積しての平均2乗誤差を示した。

4. 考察

4-1. 空間部と色彩部の合成

上では空間部と色彩部を別々のニューラルネットワークで構成したが、それらは線形であるので両者を合成することができる。

一次元カラー入力画像を行列Pで表そう。

$$P = \begin{bmatrix} r_{10} & r_{11} & \dots & r_{16} \\ g_{10} & g_{11} & \dots & g_{16} \\ b_{10} & b_{11} & \dots & b_{16} \end{bmatrix} = \begin{bmatrix} \mathbf{r} \\ \mathbf{g} \\ \mathbf{b} \end{bmatrix}$$

入力に対する中間層素子の空間重みとRGB重みをそれぞれ行列S, Cで表す。

$$S = \begin{bmatrix} s_0^1 & s_1^1 & \dots & s_{16}^1 \\ s_0^2 & s_1^2 & \dots & s_{16}^2 \\ s_0^3 & s_1^3 & \dots & s_{16}^3 \\ \vdots & \vdots & \ddots & \vdots \\ s_0^k & s_1^k & \dots & s_{16}^k \end{bmatrix} = \begin{bmatrix} \mathbf{s}^1 \\ \mathbf{s}^2 \\ \mathbf{s}^3 \\ \vdots \\ \mathbf{s}^k \end{bmatrix}$$

$$C = \begin{bmatrix} c_1^1 & c_2^1 & c_3^1 \\ c_1^2 & c_2^2 & c_3^2 \\ c_1^3 & c_2^3 & c_3^3 \end{bmatrix} = \begin{bmatrix} \mathbf{c}^1 \\ \mathbf{c}^2 \\ \mathbf{c}^3 \end{bmatrix}$$

このとき、中間層の各素子は行列

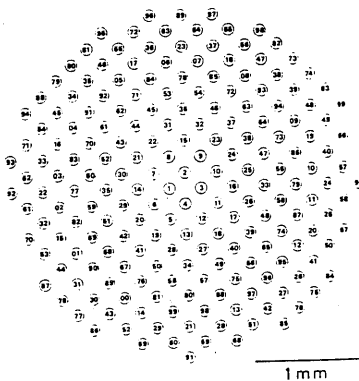
$$C [SP^T - T]^T - H = CPS^T - R$$

で表せよう。ここで、T, Hは閾値列ベクトル、 $R = CT^T - H$ とした。このように中間層の各素子の応答は入力画像に対し色彩および空間的に重み付けしたものとなる。

4-2. 網膜神経細胞との対応

発光ダイオードを199個配列し、それを実験動物コイの網膜上、図2Aのように投影した⁽²⁾。発光ダイオードを中心部から順次点滅を繰り返したときの双極細胞(OFF中心型)、水平細胞、アマクリン細胞(持続型および過渡型)の電位応答を図2Bに示す⁽³⁾。双極細胞、水平細胞、持続型アマクリン細胞は光照射中応答が持続するのに対し、過渡型アマクリン細胞は、光パルスの立ち上りや立ち下り時にインパルス応答を伴う過渡的応答を示す。網膜上に小さなスポット光を走査させたときの双極細胞(ON中心型、OFF中心型)、水平細胞、アマクリン細胞の応答を図3に示す⁽³⁾。図2および3から、双極細胞の受容野は他の細胞の受容野よりかなり狭いように見えるが、これは小さな面積のスポット光には、双極細胞の受容野周辺部が応答しないためである。円形スポット光の直径を変えた場合および、大きな円環光の内径を変えた場合のOFF中心型双極細胞応答を図4に示す⁽⁴⁾。

A.



B.

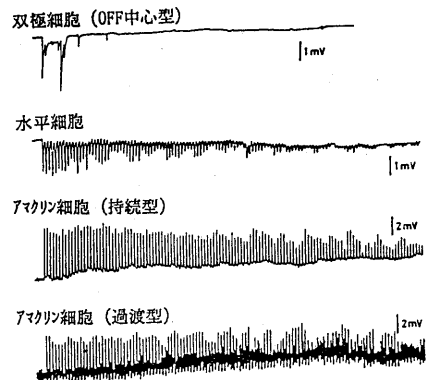


図2. A, 網膜面上でのLEDディスプレイ像⁽²⁾。各LED素子像には便宜的に、点滅の順に従い中心から周辺に向かって渦巻状に1から199の番号を付した。但し、百位は省略した。B, Aの各LED素子を一個づつ点滅したときの網膜内神経細胞の応答電位⁽³⁾。

これら網膜神経細胞の受容野特性と、図1に示したニューラルネットワークの中間層素子の空間重みを比較すると、 s_1 が水平細胞や持続型アマクリン細胞に類似し、 s_2 が過渡型アマクリン細胞に、 s_3 が双極細胞にそれぞれ対応していることが分かる。さらに色彩重みを考慮すると、 c^1 、 c^2 、 c^3 と s_1 との組合せが、水平細胞のそれぞれL型、RG型およびRGB型に対応し⁽⁵⁾、3.でも述べたが、 c^1 または c^3 と s_3 の組合せが2重対比型双極細胞⁽⁶⁾、 c^2 と s_3 の組合せが空間対比(1重対比)型双極細胞に対応する。また、過渡型アマクリン細胞は色対比を示さないで、 c^2 と s_2 の組合せがこれに対応する。その他の組合せに対応した網膜神経細胞も発見の可能性が残されている。

5. まとめ

いわゆる砂時計型をした写像型ニューラルネットワークによるカラー画像の情報圧縮について述べた。本研究で用いたニューラルネットワ

ークは線形で、また、局所空間部と色彩部に分割して考えた。色彩部は恒等写像型ニューラルネットワークで構成した。最近多く指摘されているように、恒等写像型ニューラルネットワークの中間層素子は入力データの主成分分析に類似した動作をする⁽⁷⁾。これは、網膜水平細胞が色彩入力データの主成分分析に類似した動作をしていることにも対応している^{(8)・(9)}。一方、ニューラルネットワークの局所空間部では入力素子と中間層素子間の結合重みを両者間の距離に依存して減衰するようにした。ここでの中間層素子もまた網膜神経細胞の受容野に良い対応がみられた。これら網膜神経細胞も局所空間領域における入力データの主成分分析に類似した動作をしているものと推定される。また、ニューラルネットワークの中間層の素子間に結合重みの直交性がみられるが、神経細胞にも同様の性質があり興味深い。

ここで用いたニューラルネットワークのカラー画像圧縮方式では、出力再生画像における希

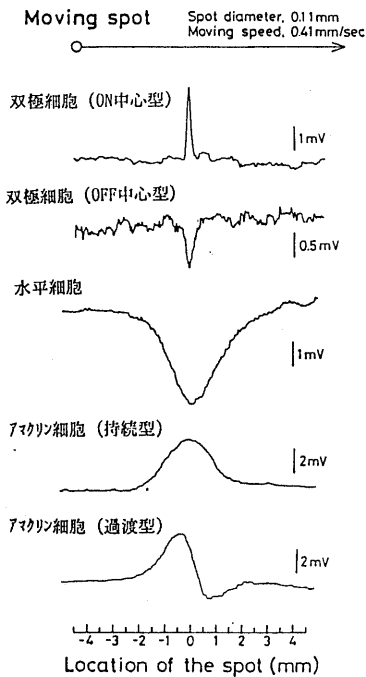


図3. 運動スポット光に対する網膜内神経細胞の応答電位⁽³⁾.

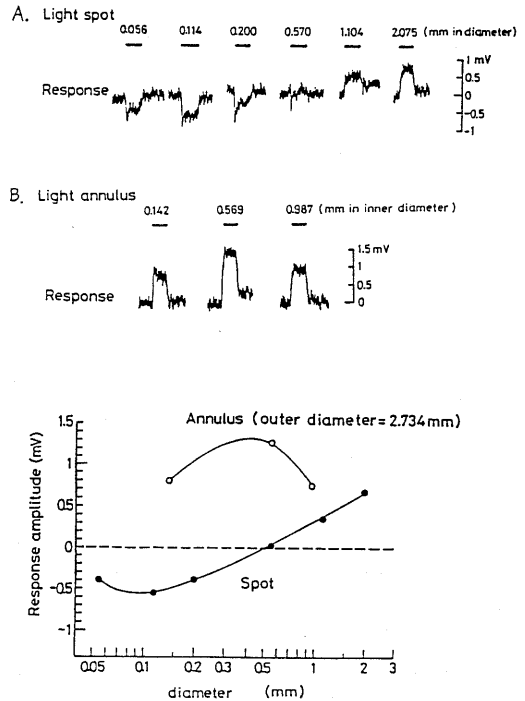


図4. 静止したスポット光の直径および円環光の内径を変えたときのOFF中心型双極細胞の応答電位⁽⁴⁾.

望の空間精度，希望の色彩精度に応じてニューラルネットワークの中間層の素子数を調整できる利点をもつ。

本研究では線形なニューラルネットワークを用いたが，必要に応じ各素子に非線形性を持たせること，さらに最終目標として，局所出力画像をいかに再構成し全体画像を再生するかが今後の課題である。

文 献

- (1) G.W. Cottrell, P. Munro, and D. Zipser: "Image compression by back propagation", Tech. Rep. 8702 UCSD, Inst. Cognitive Sci.(1987).
- (2) 袋谷賢吉: "網膜のL型水平細胞応答の空間減衰特性", 信学論(A) J61-A, 10, pp.843-850(1981-10).
- (3) 袋谷賢吉, 吉田順作: "網膜における画像の動き検出機構について", 第10回画像工学コンファレンス論文集, 8-19, pp.217-220 (1979-11).
- (4) 袋谷賢吉, 吉田順作: "網膜の双極細胞における図形情報処理について", 信学技法 ED79-32, 79, 46, pp.1-8(1979-6).
- (5) 袋谷賢吉, 長谷博行, 原 健一: "網膜における神経回路網——視細胞(錐体)と水平細胞との相互作用", 信学論(C), J60-C, 8, pp.445-452(1977-08).
- (6) A. Kaneko and M. Tachibana: "Retinal bipolar cells with double color-opponent receptive fields", Nature, 239, pp.220-221(1981).
- (7) 星野俊仁, 米倉達広, 松本哲也, 鳥脇純一郎: "3層恒等写像ニューラルネットワークによる主成分分析法の検討", 信学技法, PRU90-55, 90, 223, pp.7-14(1990-09).
- (8) K.Fukurotani: "Color information coding of horizontal cell responses in fish retina", Color Res. Appl., 7, 2 pt.2, pp.146-148(1982).
- (9) 袋谷賢吉: "網膜色覚神経回路における環境適応的波長コーディング", 信学論(D-II) J73-D-II, 8, pp.1233-1241(1990-08)