

相互情報量による最適受容野数の決定

Decision of Number of Receptive Field by Mutual Information

加藤 哲彦†
Tetsuhiko Kato松葉 育雄‡
Ikuo Matsuba

1. はじめに

ヒトの脳内における視覚情報の処理方法ということについて、いまだに解明されていない点が多い。生理学的知見からのアプローチもあるが、近年、計算機シミュレーションからのアプローチによってその仕組みを解明しようという研究が進んでいる。

ヒトが物を見ると、まず網膜の細胞が発火し、その信号が脳の後頭葉にある第一次視覚野に送られる。第一次視覚野の細胞は、視野中のある範囲からの入力を受け、この範囲を受容野という。細胞の受容野に、受容野ごとの特定の角度を持った線分が提示されるとその細胞が強く反応する。これを細胞の方位選択性という。

生理学的知見データでは受容野の方位選択性の種類は何種類が存在するかということは判明していない。判明すれば、ヒトが外界の画像を処理して、認識するとき当たって、何種類の方位があれば必要十分に外界を認識できるということを議論できる。方位選択性を説明する数理モデルとして、再構成学習のモデル[1]、競合学習のモデル[2]、相互情報量のモデル[3]などがある。

本論文では方位選択性を持った細胞の受容野を再現する数理モデルとして、相互情報量を用いた Okajima のモデル[4]を用い、細胞数をパラメータ M とし、その相互情報量を評価することによって、最適な受容野の数を決定し、それについて考察する。

2. モデルの概要

今回 Okajima のモデルを、方位選択性を再現するモデルとして採用した。このモデルは入力画像とモデル化した細胞の出力の間の相互情報量を大きくする方向に細胞間の結合 W を成長させる。単純型細胞の出力 S は

$$S_i = \sum_{x,y} W_i(x,y) f(x,y) \quad i=1,2 \quad (1)$$

複雑型細胞の出力 C は

$$C = S_1^2 + S_2^2 \quad (2)$$

入力画像は多少の位置ずれを許可し、それらを一つのクラス γ とする。相互情報量は細胞の出力 C から入力画像のクラス γ との間の相互情報量を用いる。

$$MI[C;\Gamma] = H[C] - \langle H[C|\gamma] \rangle_\gamma \quad (3)$$

一般に相互情報量を計算するのは容易でないので、入力画像の SN 比が小さい極限を仮定して計算を行う。

$$MI \propto (\langle \bar{C}_\gamma^2 \rangle_\gamma - \bar{C}^2) / (2\sigma_n^4) + B[W_1, W_2] \quad (4)$$

ここで \bar{C} と \bar{C}_γ はそれぞれ C の、平均と、クラス中の平

平均、 σ_n はノイズの分散 $\langle \cdot \rangle$ は平均操作、 B は二つの細胞間結合を正規化する働きを持つ。総合して、学習側は

$$\Delta W_i(x,y) \propto (\bar{C}_\gamma - \bar{C}) S_i f_{\gamma,m}(x,y) + \partial B / \partial W_i \quad i=1,2$$

$$\text{ただし、} (\|W_1\|^2 + \|W_2\|^2 = 1) \quad (5)$$

となる。このモジュールが一つの細胞となり、これをパラメータとして、いくつか並べてシミュレーションするのだが、同じ細胞間結合に成長しないように抑制性結合を導入する。

$$C_m \leftarrow C_m - \beta \sum_{m' < m} V_{m,m'} (C_{m'} - \bar{C}_{m'}) \quad (6)$$

V の学習側は $\Delta V_{m,m'} \propto (C_m - \bar{C}_m)(C_{m'} - \bar{C}_{m'})$ (7)
 m は細胞の番号で大きい番号の細胞が小さい番号の細胞と相関が小さくなるような抑制がかかる。 $1 \leq m \leq M$

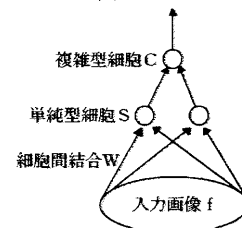


図1 モデル化された細胞

3. 有効な受容野数

入力画像は、自然画像(図 2(a))に D o G フィルタをかけたもの(図 2(b))から 32×32 pixel をランダムにサンプリングして与える。

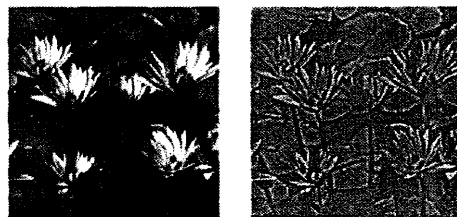


図 2 (a) 自然画像

図 2 (b) 処理画像

各細胞の受容野における入力画像と出力の間の相互情報量を式(4)を用いて算出する。ここで B の項に関しては二つの細胞間結合が正規直交するときに 0 に等しい最大値をとるので考慮しない。

抑制性結合は、細胞番号を $1, 2, \dots$ とした時にできるだけ以前の番号との相関が無くなるような抑制されるので細胞数が大きくなるほど大きい番号の受容野は成長しにくくなるが、本論文では更新回数を 64000 回に固定する。

受容野あたりの相互情報量 MI_{av} は

$$MI_{av} = 1/M \sum_{m=1}^M MI[C;\Gamma] \quad (8)$$

†千葉大学大学院自然科学研究科知能情報工学専攻

‡千葉大学工学部情報画像工学科

入力画像は 1 回の更新ごとに異なるので MI_{av} をさらに更新回数 60000 回から 64000 回の平均を取り、評価に使用する。本来、形成された受容野の相互情報量のみ抽出して平均をとり、評価すべきだが、簡単のため、近似的に評価を行う。各パラメータにおいて 5 回の試行を行う。

ここで近似のために各試行の有効受容野数 M_{eff} を導入する。有効受容野数 M_{eff} とは、更新の後、形成された受容野を目で確認し、形成されたと認められる受容野は +1、形成されていない受容野は 0 とカウントし、全て足し合わせたものをそれぞれの試行の M_{eff} とする。また、任意の一組の受容野の、方位選択性の受容野中での位置と方位が一致していた場合、同じとみなし、一方の受容野の M_{eff} も 0 とした。

下図の例では 20 個中 17 個の受容野が形成され、同じと見なせる受容野が一組あるので、 M_{eff} は 16 となった。

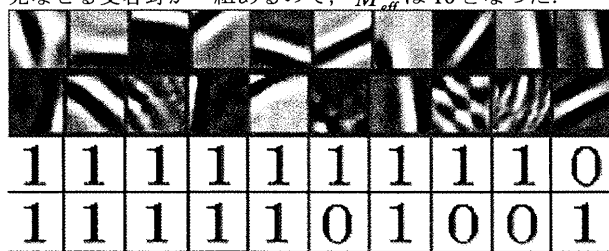


図 3 有効受容野数の例

4. シミュレーション

パラメータ M は 1, 5, 10, 15, 20, 25, 30 個を選んだ。初めに各パラメータにおける受容野あたりの相互情報量は以下ようになった。

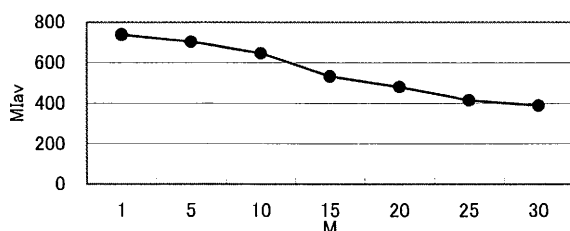


図 4 各パラメータの MI_{av} (受容野あたりの相互情報量)

次に各パラメータにおける有効受容野数を以下に示す。

表 1 有効受容野数 (上段: M 下段: M_{eff})

1	5	10	15	20	25	30
1	5	9-10	13-15	16-17	15-17	15-20

MI_{av} (受容野あたりの相互情報量) と M_{eff} (有効受容野数) をかけた値 $G = MI_{av} \times M_{eff}$ を以下に示す。

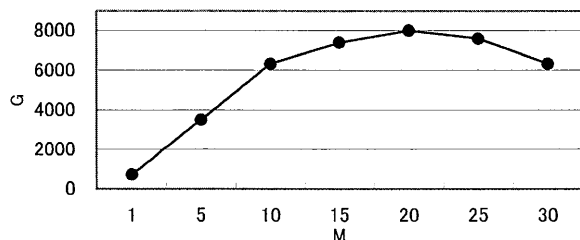


図 5 $G=MI_{av} \times M_{eff}$

図 5 は、各パラメータにおける、完全に形成されている受容野を通した、入力画像と細胞の出力の間の相互情報量の総量を表している。

細胞数が多くなるほど大きな番号を持った細胞の成長に抑制がかかるために、完全に方位選択性を持った受容野が形成される細胞数は少なくなり、全ての受容野あたりの平均を取った相互情報量 MI_{av} は小さくなっている。

受容野数 1 から 20 までは評価値 G の傾きが、平均の相互情報量の負の傾きを打ち消しながら傾きが小さくなりながらも、総相互情報量は大きくなっていくが、受容野数が 25, 30 になると、評価値 G が小さくなってしまい、受容野数 20 をピークにして総相互情報量が小さくなってしまふ。

5. 考察・結論

本論文では、視覚神経細胞の受容野を再現する数理モデルで、入力画像と細胞出力との間の相互情報量を最大化する Okajima のモデルを用いて、その相互情報量を評価することにより、このモデルにおける最適な受容野の数を決定した。

受容野数をパラメータとして、相互情報量を評価した結果、本論文での最適受容野数は 16-17 個と決定する。

単純に考えれば受容野の数が多ければ多いほど入力画像と細胞出力の間の総相互情報量は大きくなると考えられるが、そういった結果にはならなかった。このことによって、ヒトの視覚に関する脳細胞の数もただ多ければよいというのではなく、ある最適な値が存在していて、ヒトの進化の過程からヒトは現在ではその最適な数の視覚に関する脳細胞を持っているのではないかという考えに至ることができる。

本文では入力画像を自然画像としたが、構造的に単純な同心円画像を入力としたときに、最適受容野数はパラメータ $M=20$ 個のときに、12-15 個となり、上記の自然画像入力するときよりもすこし減った。このことはヒトの最適受容野数というものは、画像の複雑さにより、その画像を認識するために必要な数だけ増減するという事も考えられる。

更新回数の固定や、有効受容野数の決定方法については今後の課題としたい。

参考文献

- [1]Olshausen, B. A. ,& Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code from natural images. *Nature*, 381, 607-609.
- [2]前川聡, 喜多一, 西川禎一, 澤井秀文. (1998) 受容野の自己組織的形成. 電子情報通信学会論文誌 D-II vol.J81-D-II No.9 pp.2181-2190.
- [3]Okajima, K. (1998). Two-dimensional Gabor-type RF as derived by mutual information maximization. *Neural Networks*, 11, 441-447.
- [4]Okajima, K. (2001). An Infomax-based learning rule that generates cells similar to visual cortical neurons. *Neural Networks*, 14, 1173-1180.