

ネオコグニトロンにおけるエッジ抽出法の検討

牧野 裕樹[†] 菊池 眞之^{††} 福島 邦彦^{†††} 林 勲^{††††} 庄野 逸^{††††}

^{†, ††} 東京工科大学 コンピュータサイエンス学部 東京都八王子市

^{†††, ††††} 関西大学 総合情報学部 大阪府高槻市

^{†††††} 電気通信大学 東京都調布市

E-mail: [†]iris_3kinds@hotmail.com, ^{††}kikuchi@cs.teu.ac.jp, ^{†††}fukushima@m.ieice.org,

^{†††††}ihaya@cbii.kutc.kansai-u.ac.jp, ^{†††††}shouno@ice.uec.ac.jp

あらまし ネオコグニトロンとは視覚系のパターン認識能力を持つ階層型の神経回路モデルである。S細胞と呼ぶ特徴抽出細胞でエッジを抽出する際に、本来エッジが存在しない箇所に偽の出力を出してしまうことがある。このような偽のエッジの発生が、認識率低下の原因の1つとなっている可能性も考えられている。従来のネオコグニトロンは偽のエッジに対してC細胞の周辺抑制を導入する事により問題を解決してきた。それに対して本研究ではエッジ抽出法を改良し、偽のエッジが発生しないようにすることで、認識率の向上を図った。その結果、偽のエッジは発生しなくなったが、認識率に差は見られなかった。

キーワード ネオコグニトロン, 神経回路, S細胞, 特徴抽出

Edge Extraction for the Neocognitron

Yuki MAKINO[†], Masayuki KIKUCHI^{††}, Kunihiro FUKUSHIMA^{†††}, Isao HAYASHI^{††††}, and

Hayaru SHOONO^{†††††}

^{†, ††} Tokyo University of Technology Hachioji, Tokyo, Japan

^{†††, ††††} Faculty of Informatics, Kansai University Takatsuki, Osaka, Japan

^{†††††} University of Electro-Communications Chofu, Tokyo, Japan

E-mail: [†]iris_3kinds@hotmail.com, ^{††}kikuchi@cs.teu.ac.jp, ^{†††}fukushima@m.ieice.org,

^{†††††}ihaya@cbii.kutc.kansai-u.ac.jp, ^{†††††}shouno@ice.uec.ac.jp

Abstract Neural network model *neocognitron* has an ability of robust visual pattern recognition. Feature-extracting cells, called S-cells, in the first stage of the neocognitron, extract oriented edges from input patterns. Under certain conditions, however, the S-cells yield spurious outputs in the places where edges do not actually exist. Occurrence of the spurious edge is might lower the recognition rate. To suppress spurious responses, in the conventional neocognitron, an inhibitory surround is introduced in the input connections to C-cells, which make a blurring operation. We proposed a new method of edge extraction that suppress spurious edges, and tested how the recognition rate of the neocognitron is improved. As a result, spurious responses have been suppressed, but there was not a large difference in the recognition rate.

Key words neocognitron, neural network, S-cell, feature extraction

1. はじめに

ネオコグニトロンとは福島が提唱した、視覚パターン認識機構の神経回路モデルである [1]。予め様々なパターンを学習させておくことにより、まだ学習させていないパターンでも過去に学習したパターンと類似していれば正しく認識することが出来る。また、学習したパターンに対して変形や位置ずれ、大き

さ変化が施されたパターンも認識することが可能である。現在ネオコグニトロンは様々な研究で利用され、また、回路構造や学習方式に改良が加えられた結果、ETL 文字データベース [2] と呼ばれる手書きパターンの認識において認識率 98.6% を得ている [3] [4]。

ネオコグニトロンは S 細胞と C 細胞の階層構造を複数持つ構造を有する。神経生理学的に、S 細胞は特定の軸に沿って

接するオン領域とオフ領域からなる受容野を持つ単純型細胞 (simple cell) に似た反応特性を持つ細胞である。C 細胞は複雑型細胞 (complex cell) に似た反応特性を持ち、その前段の S 細胞が抽出した特徴の位置ずれを許容する働きをしている。この操作は、S 細胞の反応を空間的にぼかす操作であると解釈することもできる。



図 1 偽のエッジと正常なエッジ。

S 細胞が特徴抽出、C 細胞がぼかしを担う。階層の各段の C 細胞でぼかしが行われるため階層が上がるにつれ反応がぼかされ、隣接した細胞は類似した反応を示すようになる。そこで細胞数を少なくするため、階層が上がるごとに細胞配置の密度を低くしていき、最終的に 1 つの細胞が認識対象を覆うことになる。S 細胞の特徴抽出には閾値があり、高ければ細かい特徴に対する選択性が鋭くなるが変形に弱く、低ければ変形に強いが細かい特徴に対する選択性が鈍くなるという特徴を持つ。

入力の特徴抽出をする際、16 方位に対応する抽出細胞の出力をエッジと呼ぶ。図 1 の中央、左図は 22.5° のエッジ反応である。S 細胞の閾値を下げすぎると図 1 の右図のように本来存在しない位置に偽の輪郭が発生する。これを本稿では偽のエッジと呼び、認識率低下の原因である可能性がある。認識率を向上させるには、偽のエッジを生じさせない程度に閾値を上げるよりは、閾値を低く設定して偽のエッジが多少発生することを許容した上で、C 細胞の入力結合の周辺抑制の広がりがある程度大きくする。つまり S 細胞層で生じた偽のエッジを、C 細胞層で消去するほうが好ましいと考えられている。それに対してエッジ抽出法を改良し偽のエッジを発生させないようにすることで、認識率を向上を図った。

2. モデル

2.1 ネオコグニトロン

本研究では U_{S1} 層を扱う。ネオコグニトロンの各処理は数式で表すことができる。

入力層 U_{S0} の出力を $u_0(\mathbf{n})$ とする。この \mathbf{n} は、その細胞の受容野の中心位置を示す 2 次元座標である。コントラスト抽出層 U_G の受容野の中心を \mathbf{n} とすると式 (1) が与えられる。

$$u_G(\mathbf{n}, k) = \varphi \left[(-1)^k \sum_{|\boldsymbol{\nu}| < A_G} a_G(\boldsymbol{\nu}) \cdot u_0(\mathbf{n} + \boldsymbol{\nu}) \right] \quad (1)$$

k の値は 1 か 2 を取る。ここで、 $\varphi[\cdot]$ の関数は $\varphi[x] = \max(x, 0)$ によって定義される。 $a_G(\xi)$ のパラメータは、入力面から U_G 層への結合の強さを表し、釣鐘型のガウス関数の微分であるメキシカンハット型フィルターの形を取る。 U_G 層は二つの細胞面を持ち、それぞれ、on-center ($k = 2$) と off-center ($k = 1$)

型の受容野を持っている。各細胞は、入力面からの半径 A_G の範囲の重み付きの総和を求めている。単一の U_G 層への入力接続は式 (2) のようにその合計がゼロになるように設計されている。

$$\sum_{|\boldsymbol{\nu}| < A_G} a_G(\boldsymbol{\nu}) = 0 \quad (2)$$

1 段目の S 細胞層を U_{S1} 、C 細胞層を U_{C1} とし、これらの細胞層に含まれる細胞面の数をそれぞれ K_{S1} 、 K_{C1} とする。また、 U_{S1} 層の k 番目の細胞面に含まれる受容野中心位置が \mathbf{n} 、S 細胞の出力を $u_{S1}(\mathbf{n}, k)$ と記す。

S 細胞は、前段の C 細胞から興奮性結合と抑制性結合の信号を受け取っている。S 細胞層には、補助的な働きをする V 細胞が含まれている。V 細胞は抑制性細胞で、望ましくないパターンの特徴が受容野に入力された際、抑制して反応選択性を高める。また、S 細胞の出力を抑制する役割も担っている。この U_{S1} 層の κ 番目の V 細胞の出力を $v_1(\mathbf{n}, k)$ とする。以上の記号を用いて S 細胞の出力を数式で表すと式 (3) のようになる。

$$u_{S1}(\mathbf{n}, k) = \frac{1}{1 - \theta} \cdot \varphi \left[\frac{e(\mathbf{n}, k)}{h(\mathbf{n}, k)} - \theta \right] \quad (3)$$

e は興奮性入力であり、式 (4) で表される。また、 h は抑制性入力で、式 (5) で表される。

$$e(\mathbf{n}, k) = \sum_{\kappa=1}^{K_{C1}} \sum_{|\boldsymbol{\nu}| < A_{S1}} a_{S1}(\boldsymbol{\nu}, \kappa, k) \cdot u_G(\mathbf{n} + \boldsymbol{\nu}, \kappa) \quad (4)$$

$$h(\mathbf{n}, k) = b_{S1}(k) \cdot v_1(\mathbf{n}, k) \quad (5)$$

ともに κ の値は 1 か 2 をとる。

$a_{S1}(\boldsymbol{\nu}, \kappa, k)$ (θ) は、前段の κ 番目の C 細胞 $u_{C-1}(\mathbf{n} + \boldsymbol{\nu}, \kappa)$ からの興奮性可変結合の強度である。結合の広がり A_{S1} は S 細胞にいたる興奮性可変結合の結合領域と同じである。 $b_{S1}(k)$ は V 細胞からの抑制性結合の強度で、抑制は S 細胞に対して分流的に働いている。

この式から興奮性結合が大きければ大きいほど $u_{S1}(\mathbf{n}, k)$ の値は大きく、逆に抑制性結合が大きければ大きいほど値が小さくなることが読み取れる。また、式 (5) 内の v_1 と b_{S1} は式 (6)、(7) のように表す事ができる。

$$v_1(\mathbf{n}, \kappa) = \sqrt{\sum_{\kappa=1}^{K_{C1}-1} \sum_{|\boldsymbol{\nu}| < A_{S1}} c_{S1}(\boldsymbol{\nu}) \cdot \{u_G(\mathbf{n} + \boldsymbol{\nu}, \kappa)\}^2} \quad (6)$$

$$b_{S1}(\kappa, k) = \sqrt{\sum_{\kappa=1}^{K_{C1}-1} \sum_{|\boldsymbol{\nu}| < A_{S1}} \frac{\{a_{S1}(\boldsymbol{\nu}, \kappa, k)\}^2}{c_{S1}(\boldsymbol{\nu})}} \quad (7)$$

特徴抽出の時点で偽のエッジが発生する条件を無くすためには S 細胞の特徴抽出の閾値を下げすぎない事が挙げられるが、変形に弱くなり認識率が上がりにくくなる。そこで偽のエッジを予め発生しないようにしてしまえば解決するのではないかとというのが本実験の方針である。図 2 のように認識対象と S 細胞

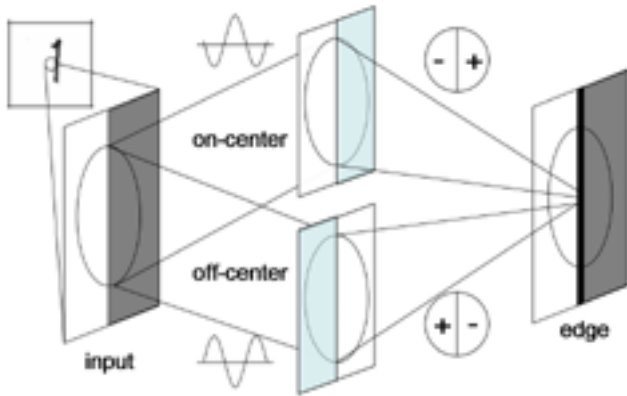


図 2 正常な反応 .

の on-center(オン中心型受容野) と off-center(オフ中心型受容野) の向きが一致しており、図の右側に太線で示されるように、それら 2 つの細胞面における反応の境目に、エッジが検出されるのが正常な反応である . on-center は受容野の中心に入力が

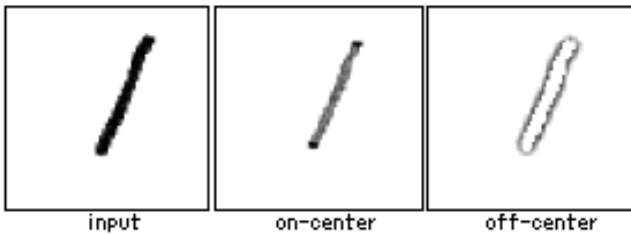


図 3 on-center と off-center .

があると刺激され、周辺に反応があると抑制される . 図 3 の中央のように input と似たような出力になるのが特徴である . 対して off-center はその逆で受容野の中心に入力があると抑制されるが周辺に反応があると刺激される . そのため input の輪郭をかたどるような出力になる . 閾値が低いと図 4 の左のような、

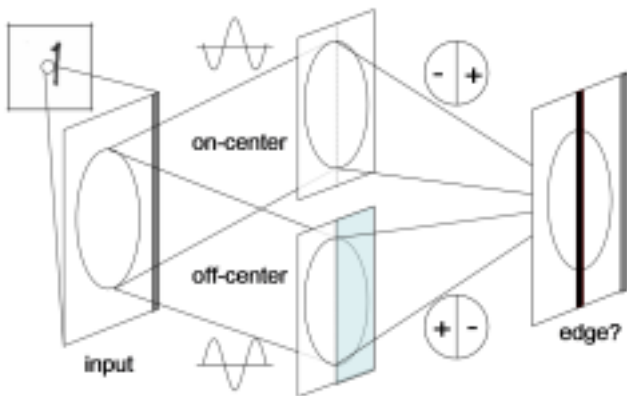


図 4 偽の反応 .

実際に輪郭がある箇所付近の、全く入力が無い位置に、方位が π だけ異なる反転した偽のエッジを出力するようになってしまう . このような偽のエッジは on-center か off-center のどちらか片方が反応するケースが多い . そこで片方のみを検出するという状況をなくすため、on-center と off-center に分けられる前に二つの積 (AND) を取り、いずれか片方しか反応しない場

合の反応を無くそうと試みた . そのためには S 細胞面の式 (3) を利用し、on-center と off-center の積をとるには式 (8) のようにする必要がある . 式内の $e(\mathbf{n}, k, \kappa)$ 、 $h(\mathbf{n}, k, \kappa)$ はそれぞれ式 (9) と式 (5) が対応する .

$$u_{S1}(\mathbf{n}, k) = \frac{1}{1 - \theta} \cdot \sqrt{\prod_{\kappa=1}^2 \phi \left[\frac{e(\mathbf{n}, k, \kappa)}{h(\mathbf{n}, k, \kappa)} - \theta \right]} \quad (8)$$

$$e(\mathbf{n}, k, \kappa) = \sum_{|\nu| < A_{S1}} a_{S1}(\nu, \kappa, k) \cdot u_G(\mathbf{n} + \nu, \kappa) \quad (9)$$

$\kappa = 1$ の時と $\kappa = 2$ の時の値の積の平方根が新たな値となる .

2.2 交差検証法

交差検証法はクロスバリデーションとも呼ばれる . 交差検証法は、データ量が少なく、テストデータの選び方によって、推定精度に大きな誤差が生じる可能性が高くなる場合や、いくつものモデルを作製したときに、どのモデルが最適であったのかを検討する際に使用する解析方法である .

まず、元のデータを同じ件数で m 個のブロックに分割する . 図 5 のように 1 つ目のブロックをテストデータ、残りのブロックをトレーニングデータとしてモデルを構築し、精度の算出を行う . それが終わったら次は 2 つ目のブロックをテストデータ、残りのブロックをトレーニングデータとしてモデルを構築し、精度の算出を行う . このような手続きを m 回繰り返すことで全てのケースがテストデータとして選ばれるため推定誤差を少なく出来る方法である .

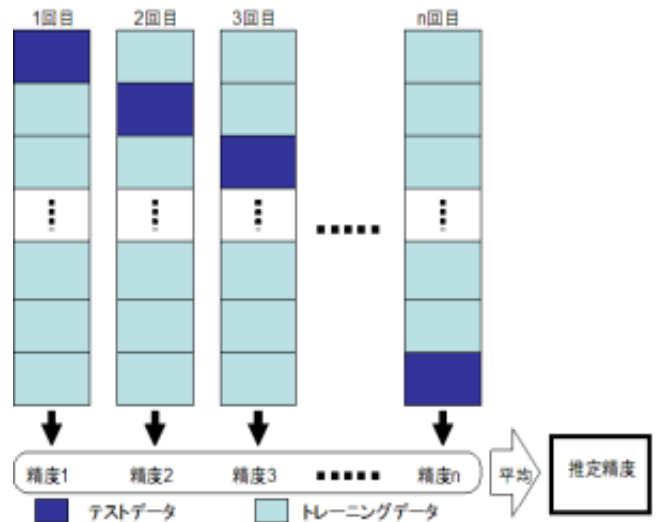


図 5 交差検証法の模式図 [5] より引用 .

交差検証法を使うと、全てのデータが一回はテストデータとして使われ、且つ、見かけ上もとのデータの $m - 1$ 倍の件数のトレーニングデータを用いてモデル構築したことになるため、少ないデータであっても推定誤差が少ない . 図 5 で m 回目に行われた学習を d^m とし、その精度を $A^{ts}(d^m)$ すると、交差検定法による推定精度 $A^{CV}(d)$ は式 (10) で与えられる [5] .

$$A^{CV}(d) = \frac{1}{m} \sum_{1}^m A^{ts}(d^m) \quad (10)$$

本研究では実験結果を統計的手法により比較するため、交差検証法をプログラムに組み込んだ。組み込んだプログラムを使用するためにブロック数の設定を作成した。また、ETLの順序に縛られないように、ランダムに再配置されたパターンの順序を作成した。それを用いた交差検証法を行う事により、学習や認識のバリエーションを増やし、出力結果の精度の向上を目指した。

3. 実験

3.1 実験条件

実験は既存のネオコグニトロンプログラムに交差検証法を実装し、S細胞面の式を式(3)にしたものを条件1、式(8)にしたものを条件2とした。1ブロック1000パターンとし、データセットのパターン数は4, 5, 6ブロック分、つまり4000, 5000, 6000個のデータを用いた。データセットのうち1ブロックをテストパターン、残りのブロックがトレーニングパターンになる。例えば4ブロック4000パターンならば1ブロック1000パターンがテストパターンに使われ、残る3ブロック3000パターンがトレーニングパターンとなる。また、データセットの0~9までの10パターンの数字を全て使用し、且つ数字の使用率が偏らないようにするため、データセット内のパターンは筆記者単位でランダムに選択した。1つのデータセットでの試行を1トライアルとし、各パターン数ごとにそれぞれ3回ずつ、条件1, 2合わせて合計18回の実験を9種類のデータセットで行った。なお2つの条件で使用しているデータセットは、パターン数別に共通の物を使用している。

3.2 結果

実験結果は表1, 2, 3のようになった。各データセットを用いて出力した認識率の平均値を出したもので、単位は%である。

表1 4000パターンの認識率

データセットの種類	条件1	条件2
4-1	97.73	97.45
4-2	98.03	97.93
4-3	98.00	98.03

表2 5000パターンの認識率

データセットの種類	条件1	条件2
5-1	97.73	97.45
5-2	98.03	97.93
5-3	98.00	98.03

表3 6000パターンの認識率

データセットの種類	条件1	条件2
6-1	97.73	97.45
6-2	98.03	97.93
6-3	98.00	98.03

条件1と条件2のそれぞれの結果の平均値について、パターン数ごとにt検定を用いたが有意差は見られなかった($p > 0.05$)。

偽のエッジが発生していないかを確認したものが図6である。数字の9と6に対してエッジ抽出を行った際の図となっている。左が改善前、右が改善後となっており、左図にうっす

らと偽のエッジが発生している事が確認できる。右図を見ると左図に発生していた偽のエッジが消えている。よって偽のエッジを取り除く試みは成功したと言える。



図6 エッジ抽出法の改善前と改善後。

4. 考察

エッジ抽出法の改良により、偽のエッジを消去する試みを行うことは成功したが、認識率の向上には直接結びつかなかった。従来のネオコグニトロンでは偽のエッジに対してC細胞の周辺抑制を導入する事により問題を解決している。つまり事前に偽の反応を取る必要が無いのではないかと考えられる。これは、何も変更していない条件1の結果と、偽のエッジを消去した条件2の結果の、どれを取っても有意差が無いためである。有意差を発生させるためにはS細胞面の閾値や、C細胞の周辺抑制のパラメータを調整し、条件1では偽のエッジが取り除けず認識率の低下が懸念されるような設定にする必要がある。

我々の脳内の方位選択細胞はガボール関数で近似されるものであり、従来型と同じく偽の反応が出ている可能性が高い。そのため、認識率を上げるためとはいえ生物学的なメカニズムを無視することは出来ないことを念頭に、まずはガボール関数を用いた実験が今後必要であると考えられる。また、本実験ではU_G層の2つの出力全てに対して積を取っていた。そのため認識に必要な情報まで除去している可能性が考えられるため、それを調べる必要がある。もし除去しているのならばどのように必要な情報を残しつつ偽の反応を除去していくかが課題となる。

文献

- [1] 福島邦彦: “位置ずれに影響されないパターン認識機構の神経回路のモデル ネオコグニトロン”, 電子通信学会論文誌 A, J62-A[10], pp. 658-665 (1979).
- [2] ETL1 database: <http://www.is.aist.go.jp/etl1cdb/>
- [3] K. Fukushima: “Neocognitron for handwritten digit recognition”, *Neurocomputing*, **51**, pp. 161-180 (2003).
- [4] Visiome Platform <http://visiome.neuroinf.jp/>
- [5] MUSASHI: “モデルの精度を推定する” <http://musashi.sourceforge.jp/tutorial/mining/xtclasssify/accuracy.html#cross>