

G-18

ネオコグニトロン型動画像識別モデルのパラメータの自動設定に対する一検討

Automatic Model Parameter Assignment for the Neo-ISRM

○大澤雅人¹, 小松崎諒², 高橋聖³, 中村英夫⁴*MasatoOsawa¹, RyoKomatsuzaki², SeiTakahashi³, HideoNakamura⁴

Abstract: We describe automatic model parameter assignment for the Neocognitron-type Image Sequence Recognition Model. In a previous procedures for learning and testing of the Model, we tried to adjust model parameters manually for good discrimination results. In this study, automatic parameter assignment is investigated by using the output signal of each category cell of the model.

1. はじめに

近年, 脳から直接情報を取得し考えるだけで機器を制御する Brain-Computer Interface(BCI)と呼ばれる技術が注目を集めている.

我々は脳機能計測装置である fNIRS (functional Near-Infrared Spectroscopy)で得られた脳活動をニューラルネットワークの Neo-ISRM(Neocognitron-type Image Sequence Recognition Model)を用いて BCI システムの構築を目指している^{[1][2]}.

先行研究では, モデルのパラメータの一つである抑制定数を設定する際に手動でその値を何度か変化させて最良の結果識別結果を求めていた.

そこで, 本研究では Neo-ISRM に抑制定数の自動設定を組み込むことで一回の識別で最良の識別結果を得られるように目指す.

2. fNIRS

fNIRS は近赤外光を用いて脳の血流変化を計測する装置である. fNIRS を用いる理由として被験者の拘束が小さくリアルタイムでの測定が可能であり, 装置が比較的小さいといった利点がある.

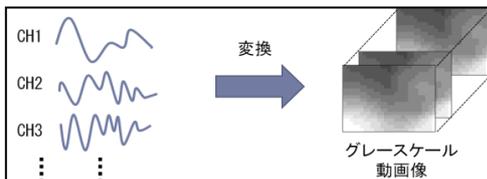


Figure1. The picture conversion of a fNIRS signal

fNIRS による測定データは Figure1 のように多チャンネルのアナログ信号として得られるため, 動画像識別モデルである Neo-ISRM によって識別を行うためにはその波形を動画像の形へと変換する必要がある.

3. Neo-ISRM

Neo-ISRM の基本構造は, パターンの特徴を抽出す

る S 細胞の層(US 層)と, 抽出された特徴の位置ずれや時空間的な伸縮を許容する働きをする C 細胞(UC 層)により構成された 2 層を 1 段とした階層構造となっており, 図 2 は 4 段構造の場合を示してある. 各層には同じ特徴に反応する 3 次元配置の細胞の集まりが複数存在し, これらをネオコグニトロン細胞面に対して細胞群と呼ぶことにする. U0 層は入力層に相当し, 動画像はあるフレーム間隔ごとに U0 層に入力される. 最終段の C 細胞層(Figure2 の UC4)はカテゴリ層であり推定するカテゴリの数だけ細胞が並ぶ.

Neo-ISRM の基本的な識別原理はネオコグニトロンと同様であり, 時系列パターンである動画像の持つ時間的伸縮やパターンの位置ずれ及び変形を許容する高い識別能力を持つ.

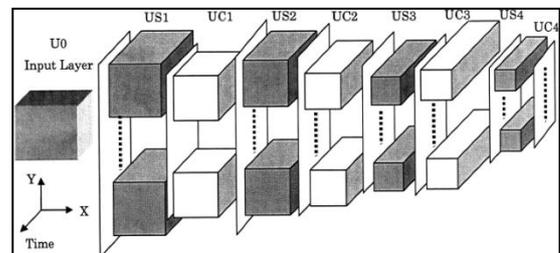


Figure2. The structure of Neo-ISRM

モデルの段数は, 識別するパターンの複雑さや解像度によって異なる. US1 層の S 細胞は, ある時刻 t における, 第 k_l 群の細胞群内の出力は次式で求められる.

$$U_{S_l}(n, k_l) = r_l \cdot \varphi \left[\frac{1 + e_{s_l}}{1 + \frac{r_l}{1 + r_l} \cdot b_l(k_l) \cdot U_{v_l}(n)} - 1 \right] \quad (1)$$

n は細胞の位置を表す 3 次元ベクトルで入力パターン U0 の座標に対応し, r_l は S 細胞に対する抑制の強さを制御するパラメータ(抑制定数)である.

抑制定数の値は類似パターンを区別する能力と, 形の変形を許容する能力との兼ね合いを考慮し設定する.

1 : 日大理工・学部・子情 2 : 日大理工・院 (前)・情報 3 : 日大理工・教員・子情 4 : 日大理工・教員・子情

4. 識別条件

本研究で, Neo-ISRМ に用いた計測データは, 20 代の言語障害のない被験者に PC の画面への文字の現示により教示し, fNIRS 装置(OMM-3000 SHIMADZU)を用いて, 左半球の前頭前野および側頭葉(全 22ch)を 115ms のサンプリング周期で計測したデータである.

rest では被験者は安静にし, Cf(Category fluency)では与えられた「種類」に分類できる単語を, Lf(Letter fluency)では与えられた文字を「頭文字」とする単語を次々にイメージしてもらう課題である.

5. パラメータの自動設定

Neo-ISRМ の抑制定数を自動で設定するために今回, 次のような評価方法を考案した.

(1) 正規化

識別結果の出力の最大値を基準にし, 最大値を 1 とし正規化する.

(2) 積分

次にそれぞれのカテゴリ細胞の出力信号を積分し, 入力パターンがカテゴリ細胞と同じ A の場合.

$$\text{Score} = \int \text{Adt} - \int \text{Bdt} \quad (2)$$

逆に入力パターンがカテゴリ細胞と同じ B である場合.

$$\text{Score} = \int \text{Bdt} - \int \text{Adt} \quad (3)$$

(3) 評価

Score が正の値であれば正しく識別, 負の値なら誤識別という判断をした.

今回入力パターンに用いた語想起は C1(Clothes), L3(KI), C2(Vegetables), L4(SI), 学習パターンには C1(Clothes), L3(KI)を用いた. Figure3 は抑制定数を二段目 10, 三段目 1 の正規化した後の識別結果であるが, 入力パターンである C1~L4 の 4 個全てが正の値であることが望ましい. この 4 個の入力パターンを積算し Score としてまとめたものが Table1 である. Table1 からわかるように入力パターンが C1 のとき, 学習パターン Cf の積算より Lf の積算のほうが大きい値となり Score が負の値で正しく識別できていないことがわかる.

Table1. The result under the paramater(A value of the second step is 10, A value of the third step is 1)

		CfAddition	LfAddition	Score
Input pattern	C1	0.29	0.59	-0.30
	L3	0	5.29	5.29
	C2	0.33	0	0.33
	L4	0.02	0.43	0.41

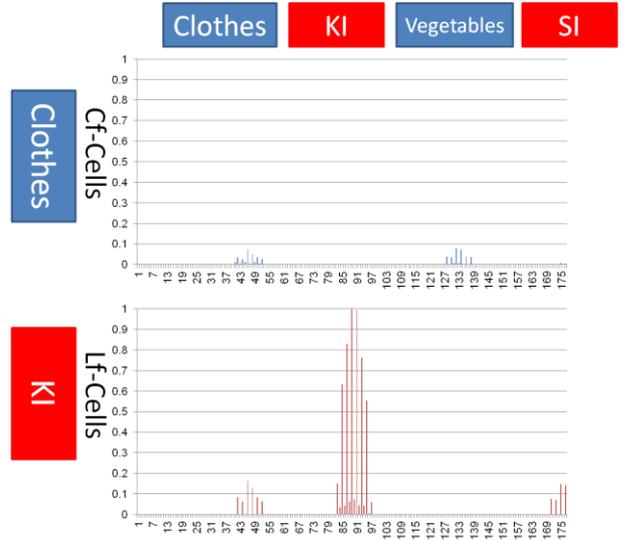


Figure3. The result normalized by 1

Figure4 は抑制定数を三段目は 1 で固定し, 二段目の抑制定数を 1 から 30 まで変化させ, 入力パターンである C1~L4 での Score をそれぞれ示したものである. 図 5 から 4 個の Score が全て正の値である 3 から 7 の抑制定数を用いれば適切な識別ができると考えられる.

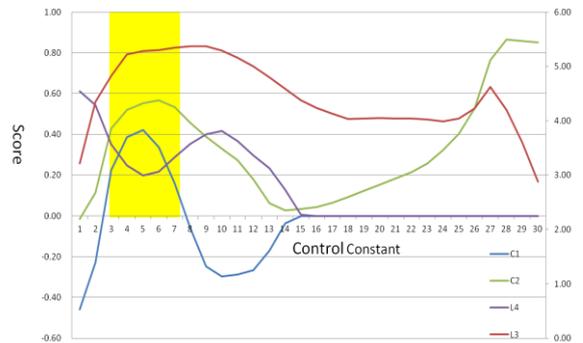


Figure4. The result when the parameter value of second step at varying 1-30

6. まとめ

今回, 抑制定数の解析方法を考案し, 最適な抑制定数の値を導き出すことができた. この解析方法を用いて自動で抑制定数を決めるプログラムを作成する. また今回の解析方法以外の方法も考案し, 比較する.

7. 参考文献

[1] 高橋聖, 横田倫一, 関根好文, 中村英夫:「ネオコグニトロン型ニューラルネットによる唇動画像からの視覚音声認識」, 電子情報通信学会, 2002 年.
 [2] 齋藤永子:「BCI 構築を目的としたネオコグニトロン型動画像識別モデルを用いた fNIRS データの識別に対する検討」, 2009 年度 修士論文.