

P300 speller における二画面インタフェースに関する検討

A Study on interface of P300 speller with second display

山本裕也 *1 吉川大弘 *2 古橋武 *2
Yuya Yamamoto Tomohiro Yoshikawa Takeshi Furuhashi

*1名古屋大学工学研究科
Graduate School of Engineering Nagoya University

Brain-Computer Interface (BCI) is a system that allows a user to control external devices and to communicate with other people by his/her thought. P300 speller is one of the BCI systems, which uses P300 as the feature to input letters. When a user inputs letters, higher accuracy and shorter input time are important for P300 speller. In this paper, we propose a new interface with second display in P300 speller that can be switched by choosing the command "next" or "back" to improve the performance of P300 speller. This paper shows that the comparison result in terms of accuracy and input time between the conventional interface and the proposed one.

1. はじめに

近年、人間の活動時に生じる脳信号をもとに、四肢を用いずにコンピュータなどの外部機器の操作を可能にする Brain-Computer Interface (BCI) の研究が盛んに行われている [Wolpaw 02][木村 11]. このような BCI は、筋委縮性側索硬化症 (ALS) 患者 [Kaub-Wittmer 03] などの、脳活動は正常であるが、体に重度の四肢麻痺を抱え意思疎通が難しい人のための、コミュニケーションツールとして発展が期待されている. BCI の一つとして、事象関連電位 (ERP) の一種である P300 を特徴量に用い、ユーザが思考のみで文字入力を行う P300 speller [Farwell 88] が存在する. P300 speller では、様々な形態のインタフェースが考案されているが、本稿においては、図 1, 図 2 のような、行列型に文字が配置されているインタフェースを対象とする. 行列型インタフェースにおいては、各行および各列がランダムに点灯することで、視覚的な刺激を使用者に呈示する. 使用者は選択したい文字に対して意識を向けることで、選択したい文字が含まれる行や列の点灯時に P300 が誘発される. この P300 を捉えることにより、使用者の選択したい文字が特定され、文字入力が可能となる. なお、本稿ではすべての行と列が 1 度ずつ点灯することを 1 シーケンスと呼ぶ.

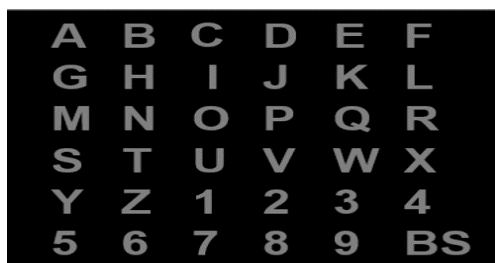


図 1: 英語入力インタフェース

Farwell らによって開発された P300 speller のインタフェースは、図 1 に示すアルファベットが表示された英字インタフェースであり、従来は日本語入力においても、このインタフェース
連絡先: 山本裕也, 名古屋大学大学院工学研究科, 名古屋市千種区不老町, 052-789-2793, 052-789-3166, yamamoto@cplx.cse.nagoya-u.ac.jp



図 2: 日本語入力インタフェース

を用いたローマ字入力が行われていた. しかし、実際の ALS 患者らによる臨床実験において、ローマ字入力では日本語を直感的に入力できないため、一文字一文字に若干の思考時間が必要であり、また文字入力の正答率が低下する結果となった. そこで、英字インタフェースに代わり、ひらがな (50 音) が 6×10 の行列型に配置された、図 2 に示す日本語入力インタフェースを用いることとする.

しかし一方でひらがな表示は、アルファベットに対して文字数が増加するため、英字のインタフェース (6×6 行列型) よりも選択肢が多く必要となる. その際、この行列サイズの増加により、1 シーケンスあたりの点灯時間が長くなる、文字間隔が狭くなるなどの要因により、入力時間の増加、正答率の低下といった問題が発生する可能性がある. そこで本稿では、これらの要因を排除することによる入力時間の短縮および正答率の向上を目的として、新しい日本語入力インタフェースの提案を行う. 提案するインタフェースでは、<次>および<戻>のコマンドを入力することにより、ひらがなの前半と後半とを切り替えて文字の入力を行う. 実験において従来インタフェースとの比較を行い、提案インタフェースにより文字の入力性能が向上することを示す.

2. 提案インタフェース

2.1 二画面インタフェース

従来の日本語入力においては、図 2 のように、1 つの画面に 50 音および記号の選択肢が表示された、 6×10 の行列型インタフェースが用いられてきた. しかし、このインタフェースでは、ひらがな表示による行列サイズの増加が原因で、いくつか

の問題点が生じる。一つは、1シーケンスあたりの点灯数の増加である。また、画面に表示される文字数が増加することで、文字同士の間隔が狭くなり、使用者が本来入力したいターゲットの周辺の文字に対して誤入力頻度が増加することが考えられる。

そこで本稿では、図3のように、1つの画面を二分割することで、1画面に表示される行列サイズを6×5と小さくし、文字間隔を広くとったインタフェースを提案する。このインタフェースでは、使用者には「あ～な行」を表示した図3の右側の画面が最初（文字の入力ごと）に表示されており、<次>や<戻>コマンドを入力することにより画面の切り替えを行うことができる。なお、図3の右側の画面を第一画面、左側の画面を第二画面とし、インタフェースに表示されている、ひらがなや記号のひとつひとつを1コマンドと呼ぶ。また本稿では、図2の従来の日本語インタフェースを1matrix、図3の提案インタフェースを2matrixと呼ぶ。

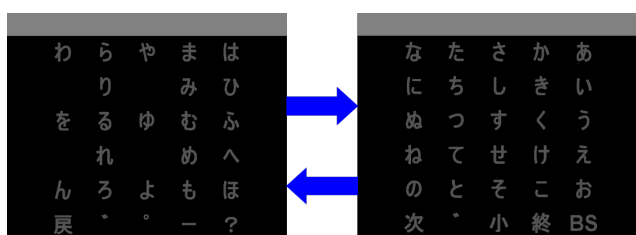


図3: 提案インタフェース

表1に、1matrixと2matrixの特徴の比較を示す。1つの画面に表示される文字数は、1matrixと比較して2matrixが少なくなるため、文字間隔が広くなり、ターゲット周辺文字への誤入力確率が低くなると期待できる。一方で、2matrixにおいては、二画面目にある文字を入力する場合には、<次>の画面移動のコマンドとターゲット文字の2度の入力を成功させる必要があるため、2コマンド分の入力時間が必要であることや、正答率が低下する（正答率が90%のとき、二画面目の文字に対する正答率は81%（0.9×0.9）となる）ことが考えられる。

表1: インタフェース間の比較

	1matrix	2matrix
文字間隔	狭い (1画面に10列)	広い (1画面に5列)
点灯数(1seqあたり)	16点灯 (6行+10列)	11点灯 (6行+5列)
行列の画面移動	不要	必要

2.2 文章入力時間の算出

Google日本語コーパスにおけるuni-gramのデータより、一般に文字の出現頻度は、「あ～な行」のある第一画面と、「は～わ行」のある第二画面で7:3と算出された。そこでここでは、この出現頻度に基づき、各インタフェースにおける文章入力時間の算出方法について述べる。

本稿では文章を完全に正しく入力した際の入力時間として、1matrixにおける入力時間を式(1)により、また2matrixにおいては式(2)により算出する [Dal Seno 09]. なお文章の入力にあたり、本稿では“BS”を用いることにより、誤判別され

てしまった文字を消去し、再度正しい文字の入力を行うことで完全な文章の入力を想定する。また、入力時間の算出にあたっては、“BS”を入力する際の時間も考慮し、式(2)については<次>や<戻>を入力する際の時間も考慮する。

$$T_1 = \frac{a_1 \cdot s \cdot n}{2p - 1} \tag{1}$$

$$\begin{aligned} T_2 &= \frac{a_2 \cdot s \cdot 0.7n}{2p - 1} + 2 \frac{a_2 \cdot s \cdot 0.3n}{2p - 1} \\ &= 1.3 \frac{a_2 \cdot s \cdot n}{2p - 1} \end{aligned} \tag{2}$$

T_1, T_2 はそれぞれのインタフェースにおける全文字の入力時間を示している。式(1)、式(2)において、 p は判別正答率、 n は入力文字数、 s は刺激呈示間隔、 a_1 は1matrixでの、 a_2 は2matrixでの平均点灯数を表している。また、式(2)における右辺第一項は第一画面に配置された文字の入力時間を、第二項は第二画面に配置された文字の入力時間を表している。第一項および第二項における0.7および0.3は、各画面に配置された文字の出現頻度を表しており、第二項の係数2は第二画面における文字の入力を行う際に、<次>と「ターゲット文字」の2つのコマンドの入力が行われることを表している。ここで式(1)および(2)から、2matrixにおける入力時間は

$$T_2 = 1.3 \cdot \frac{a_2}{a_1} \cdot T_1 \tag{3}$$

と表される。この時、どちらのインタフェースにおいても同程度のシーケンス数で文字入力が可能であると仮定すると、

$$a_1 = \frac{16}{11} \cdot a_2 \tag{4}$$

となる。よって、2matrixにおける入力時間の期待値は、

$$T_2 = 0.89T_1 \tag{5}$$

として求められる。このように、2matrixにおける入力時間は、1matrixに対して0.89倍となることが期待され、2matrixを用いることにより入力時間の短縮が可能となると考えられる。

3. 実験

本稿では、1matrixおよび2matrixのインタフェースについて、オフライン実験を行い、正答率および入力時間の比較、検討を行った。

3.1 使用データ

本実験には、予め計測した4名(Sub1～Sub4)のP300 speller使用時の脳波データを用いた。計測には、デジテックス研究所製のPolymate AP216を用い、図4に示す国際10-20法 [H.H.Jasper 58] に従った5電極(Fz, Cz, Pz, O₁, O₂)、及び基準電極としてA2(耳朶)に電極を配置し、サンプリング周波数100Hzで計測した。

刺激呈示間隔を400ms(点灯時間を200ms、消灯時間を200ms)とし、図2および図3の2種類の日本語インタフェースにより文字の入力を行った。各被験者がそれぞれのインタフェースにおいて、指定された40文字分の入力を行い、得られた脳波データのうちターゲット点灯時の脳波データには‘P300データ’、非ターゲット点灯時の脳波データには‘non-P300データ’のラベル付けを行った。この際、疲れなどによる影響を考慮して、被験者のうち2人は1matrix→2matrixの順に、残りの2人に対しては2matrix→1matrixの順にデータの計測を行った。各インタフェースにおいて、1文字の入力に対し、それぞれ10シーケンス分の点灯を行った。

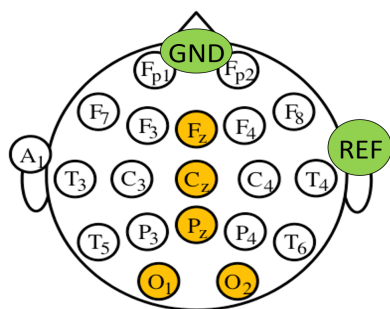


図 4: 使用電極

3.2 実験方法

本実験では、文字入力時の正答率および1文字あたりの入力時間を算出する。事前学習には、各インタフェースにおける12文字分の脳波データを用いた。1matrixにおける学習データは、P300データが240点灯分、non-P300データが1680点灯分あり、2matrixについてはP300データが240点灯分、non-P300データが1080点灯分で構成されている。すなわち、事前学習の計測時間は、1matrixの方が2matrixよりも長い。この事前学習用データに対し Stepwise Linear Discriminant Analysis (SWLDA) [Krusienski 06] を適用し、学習を行った。文字入力を行うテストセッションでは、学習に用いたもの以外の脳波データから、ターゲット点灯時には‘P300データ’を、非ターゲット点灯時には‘non-P300データ’をランダムに選択し、用いた。1試行で10文字の判別を行い、毎回異なる事前学習データおよびテストデータを用いて、これを100試行を行った。また、文字判別の手法としてはRB-ARQ[Takahashi 09]を用いた。これは、予め閾値を設定し、ある文字がターゲットであると思われる確率である“事後確率”を随時計算しながら、その中で最大の値である最大事後確率が閾値を超えるまで刺激をランダムに呈示、閾値を超えたところでターゲットとして判別を行う手法である。この手法により、点灯数を動的に決定することが可能となる。本実験においては、閾値が0.95と0.9の2通りの場合について検証した。ただし、RB-ARQにおいて、1文字あたりの最大点灯数は10シーケンスとした。

3.3 結果と考察

表2に、RB-ARQにおける閾値が0.95の場合での、4人の被験者のそれぞれのインタフェースでの正答率および入力時間の結果を示す。正答率は、1000文字の入力のうち正しく入力できた文字の割合を示したものであり、入力時間は1文字を入力する際にかかる時間の平均値を示したものである。入力時間は式(1)、式(2)において $n=1$ として算出した。表2より、それぞれの被験者について、インタフェース間で正答率に差がないことがわかる。これは、RB-ARQを用いていることで、判別率に差がある場合には、正答率ではなく判別時間の短縮という形で差が表われるためである。また、表3に、各被験者における入力時間の実験値 T_1 (1matrix)、 T_1 より算出した2matrixにおける入力時間の期待値である $0.89T_1$ 、および実験値 T_2 (2matrix)の比較を示す。表より、Sub2、Sub3については、実験値 T_2 が期待値 $0.89T_1$ よりわずかに長くなっていることがわかる。これはRB-ARQにおいて、1点灯ごとに各文字がターゲット文字である事後確率を計算する際に、1matrixでは1シーケンスあたりの点灯数が多いことで、同じ1シーケンスあたりの確率更新において多くの情報量が用いられることに起因すると考えられる。これにより、2節で述べた、どちらのインタフェースにおいても同程度のシーケンス数で文

字入力が可能であるという仮定が成立せず、理論値との差が生じたと考えられる。ただし、ほとんどの被験者で実験値 T_1 (1matrix)よりも実験値 T_2 (2matrix)が小さくなるという結果が得られた。

表 2: 正答率および入力時間 (閾値: 0.95)

	1matrix		2matrix	
	正答率	入力時間 [秒]	正答率	入力時間 [秒]
Sub1	0.94	17.5	0.95	16.0
Sub2	0.94	15.9	0.93	15.7
Sub3	0.90	50.6	0.88	50.7
Sub4	0.93	19.5	0.93	14.8

表 3: 期待値と実験値の比較 (閾値: 0.95)

	実験値 T_1 (1matrix)[秒]	理論値 $0.89T_1$ [秒]	実験値 T_2 (2matrix)[秒]
Sub1	17.5	15.6	16.0
Sub2	15.9	14.2	15.7
Sub3	50.6	45.0	50.7
Sub4	19.5	17.3	14.8

表4に、閾値が0.9の場合における、それぞれのインタフェースでの各被験者の正答率および入力時間の結果を示す。こちらについても閾値が0.95の場合と同様に、それぞれの被験者について、インタフェース間で正答率に差がないことがわかる。また、ほとんどの被験者で実験値 T_1 (1matrix)よりも実験値 T_2 (2matrix)が小さくなった。Sub3については、いずれの閾値の場合でも他の被験者と同程度の正答率を得るために、より多くの入力時間が必要であった。これはSub3においてP300データとnonP300データの差分が他の被験者よりも小さかったためであると考えられる。

表 4: 正答率および入力時間 (閾値: 0.90)

	1matrix		2matrix		$0.89T_1$ [秒]
	正答率	入力時間 [秒]	正答率	入力時間 [秒]	
Sub1	0.92	16.3	0.90	15.5	14.5
Sub2	0.88	15.6	0.88	15.5	13.9
Sub3	0.85	49.0	0.83	51.0	43.6
Sub4	0.87	19.5	0.90	14.3	17.3

上記で求めた正答率と入力時間はトレードオフの関係にあるため、この両者を同時に評価する評価指標として、Utility[Dal Seno 09]の算出を行った。Utilityは、誤判別により誤った文字が入力された際に、“BS”によりその文字を消去し、完全に正しい文章を入力する場合の情報伝達速度を表している。なお、Utilityは、“BS”を文字候補の一つとし、入力によりどのような文字の情報が得られるかを利用した相互情報量であるため、2matrixにおける<次>や<戻>は情報を持つ文字候補として扱わない。Utilityは以下の式により算出さ

れる。

$$\text{Utility} = \frac{(2p - 1) \log_2(N - 1)}{d} \quad (6)$$

ただし p は判別正答率, N は選択できる候補数 (本実験では 1matrix : $N=56$, 2matrix : $N=54$), d は 1 文字を入力するのに要した平均時間である。図 5(a) は、閾値が 0.95 の場合、図 5(b) は閾値が 0.90 の場合の、各被験者におけるそれぞれのインタフェースでの Utility を表している。図より、閾値がいずれの場合においても、ほとんどの被験者に対して、2matrix により Utility の向上が確認できた。この結果に対して、被験者間の影響を除いた検定を行ったところ、インタフェースの違いに対して Utility に統計的有意差があるという結果が得られた。(ボンフェローニにより多重比較を考慮。閾値が 0.95 の場合 : p 値 = $1.67 \times 10^{-8} < 0.025$ (= 0.05/2), 閾値が 0.90 の場合 : p 値 = $1.88 \times 10^{-5} < 0.025$ (= 0.05/2)。これにより、2matrix を用いることで、事前学習の量を削減しながらも、正答率を保ち、1 文字あたりの入力時間が 1matrix より短縮できることが確認できた。

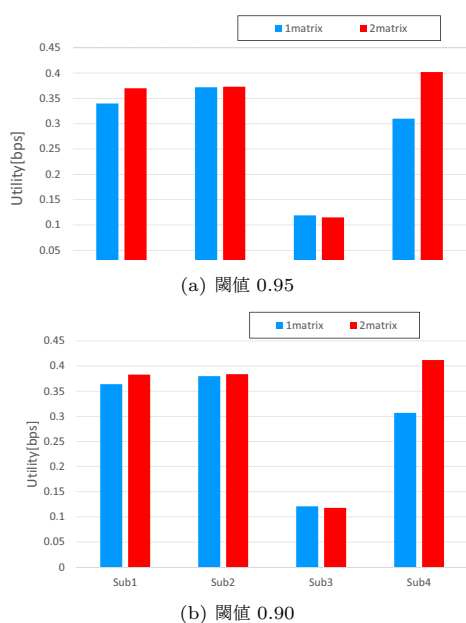


図 5: 各インタフェースにおける Utility

4. まとめ

本稿では、P300 speller において、ユーザの負担軽減を目的として、従来の日本語インタフェースを二つの画面に分割した二画面入力インタフェースを提案した。実験の結果、2matrix を用いることにより、文字入力の総合指標である Utility が向上することを示した。今後の課題として、2matrix での文字間隔や文字の大きさの違いが及ぼす、正答率への影響に関する検討、および二画面入力におけるユーザの心理的負担に対する検討が挙げられる。

参考文献

[Dal Seno 09] Dal Seno, B., Matteucci, M., and Mainardi, L.: The Utility Metric: A Novel Method

to Assess the Overall Performance of Discrete Brain-Computer Interfaces, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 18, No. 1, pp. 20–28 (2009)

[Farwell 88] Farwell, L. and Donchin, E.: Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials, *Electroencephalography and clinical Neurophysiology*, Vol. 70, No. 6, pp. 510–523 (1988)

[H.H.Jasper 58] H.H.Jasper, : The ten-twenty electrode system of the international federation, *Electroencephalography and clinical neurophysiology*, Vol. 10, No. 2, pp. 371–375 (1958)

[Kaub-Wittemer 03] Kaub-Wittemer, D., Steinbuchel, N., Wasner, M., Laier-Groeneveld, G., and Borasio, G.: Quality of life and psychosocial issues in ventilated patients with amyotrophic lateral sclerosis and their caregivers, *Journal of pain and symptom management*, Vol. 26, No. 4, pp. 890–896 (2003)

[Krusienski 06] Krusienski, D. J., Sellers, E. W., Bayoudh, S., McFarland, D. J., Vaughan, T. M., and Wolpaw, J. R.: A comparison of classification techniques for the P300 Speller, *Journal of Neural Engineering*, Vol. 3, No. 4, pp. 299–305 (2006)

[Takahashi 09] Takahashi, H., Yoshikawa, T., and Furuhashi, T.: A study on application of reliability based automatic repeat request to brain computer interfaces, in *Advances in Neuro-Information Processing*, pp. 1013–1020, Springer (2009)

[Wolpaw 02] Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M.: Brain-Computer interfaces for communication and control, *Clinical Neurophysiology*, Vol. 113, No. 6, pp. 767–791 (2002)

[木村 11] 木村達洋, 山崎清之: 最近の Brain-Computer-Interface (BCI) 研究の動向, 東海大学紀要. 開発工学部, Vol. 20, pp. 7–12 (2011)