

RB-ARQにおけるN-gram情報の利用とP300 spellerへの適用

佐溝 恵里[†] 吉川 大弘[†] 古橋 武[†]

[†]名古屋大学 大学院工学研究科 計算理工学専攻 〒464-8603 愛知県名古屋市千種区不老町
 E-mail: [†]samizo@cmlpx.cse.nagoya-u.ac.jp, ^{††}{yoshikawa, furuhashi}@cse.nagoya-u.ac.jp

あらまし P300 speller は、脳波から得られる P300 を特徴量に用い、ユーザーが思考によって文字入力を行う Brain Computer Interface (BCI) の一種である。ここで用いられる P300 は SN 比が悪く、判別精度向上のためには加算平均を行う必要があるが、これによる入力速度の低下が問題となっている。一方、文章の入力においては、ある文字の出現確率やある文字の次に生じる文字の起こりやすさ、すなわち文字間の遷移確率の偏りが存在する。そこで本稿では、日本語 P300 speller において、ひらがなの出現確率や文字間の遷移確率を事前確率により考慮する手法を提案し、従来手法と比較して、P300 speller における入力速度が向上することを示す。

キーワード BCI, P300 speller, N-gram, 文字間遷移確率, 事前確率

Application of RB-ARQ using N-gram information to P300 speller

Eri SAMIZO[†], Tomohiro TOSHIKAWA[†], and Takeshi FURUHASHI[†]

[†] Dept. of Computational Science and Engineering, Graduate School of Engineering, Nagoya University,
 Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan

E-mail: [†]samizo@cmlpx.cse.nagoya-u.ac.jp, ^{††}{yoshikawa, furuhashi}@cse.nagoya-u.ac.jp

Abstract P300 speller is one of the Brain Computer Interfaces, which uses P300 as the feature quantity and allows users to select letters just by thinking. Because of the low signal-to-noise ratio of the P300, signal averaging is often performed to improve the spelling accuracy instead of the degradation of the spelling speed. In texts, there is variability in occurrence probabilities and transition probabilities between letters. This paper proposes Japanese interface of P300 speller considering the occurrence probabilities and the transition probabilities as the priori probabilities. It shows that the spelling speed is improved by the proposed method comparing with the conventional method.

Key words BCI, P300 speller, N-gram, transition probability between letters, priori probability

1. はじめに

Brain Computer Interface (BCI) は、脳活動を測定、解析し、その情報を基に思考判別を行い、外部機器の操作を試みるインターフェースであり [1]、筋萎縮性側索硬化症 (ALS) 患者のような重度の運動障害を抱えた人に対するコミュニケーションツールとしてその発展を期待されている。脳活動の計測では、非襲性・安価であり、最も BCI として現実的な Electroencephalogram (EEG) が用いられることが多く、Farwell らが開発した P300 speller [2] は、EEG から得られる事象関連電位 (ERP) の一種である P300 を特徴量として用いる BCI のコミュニケーションツールである。P300 speller では、主に文字が行列型に配置されたインターフェースを用いて、ユーザーの意図する文字を含む行および列の点灯によって P300 を誘発させることで、文字入力を可能にする。

しかし、ここで特徴量として用いられている P300 は、一般的に SN 比が悪いため、加算平均を行うことで判別精度を向上させている [3] [4]。しかしこれは判別時間の増加につながり、実用面を考えた時、ユーザーの負担を軽減するためには、短い判別時間で正確に入力可能であることが望ましい。これに対し、固定した刺激数を呈示する従来の刺激呈示法に対して、動的に刺激数を決定する刺激呈示法である信頼度に基づく自動再送要求 (Reliability-Based Automatic Repeat reQuest: RB-ARQ) が提案され [5]、P300 speller で、判別時間の増加を抑えつつ、判別正答率の改善が行えることが報告されている [6]。

RB-ARQ では、各文字について、点灯の前にその文字がターゲットとなりうる確率、すなわち事前確率が等確率で割り振られている。一方、文章においては、ある文字の出現確率の偏りや、ある文字の次に生じる文字の起こりやすさ、すなわち文字間の遷移確率の偏りが存在し、文章理解や音声認識の分野では、

この文字間のつながりを考慮した文字訂正や入力・認識補助などが広く用いられている [7] [8] .

そこで本稿では、日本語入力を目的とした日本語インターフェースの P300 speller において、RB-ARQ の事前確率に各ひらがなの出現確率やひらがな文字間の遷移確率を用い、前の入力文字に応じて、入力確率の高い文字の事前確率を高める方法を提案する。3名の被験者による日本語 P300 speller での文字入力実験を行い、従来手法と比較して判別時間の短縮が可能となることを示す。

2. RB-ARQ

RB-ARQ とは、それぞれの文字がターゲットであると思われる確率（事後確率）を随時計算しながら、その中で最大となる最大事後確率があらかじめ設定した閾値を超えるまで刺激をランダムに呈示し、その後、最大事後確率を持つ文字をターゲットとして判別する手法である [6]。時間 t における EEG データから抽出された特徴ベクトルを \mathbf{x}_t とし、時刻 T における EEG データの集合を $X_T = \{\mathbf{x}_t | t = 1, 2, \dots, T\}$ とすると、事後確率は以下の式で表せる。

$$P(k|X_T) = \frac{P(k) \prod_t p(\mathbf{x}_t|k)}{\sum_{l \in K} P(l) \prod_t p(\mathbf{x}_t|l)} \quad (1)$$

ここで、 K はとりうるラベル（文字）の集合であり、 $k \in K$ となる。また、 $P(k)$ はデータがラベル k に属する事前確率（判別開始前にターゲットとなりうる確率）で、点灯開始前は全ラベルに対して等しく設定されている。事後確率は、事前確率 $P(k)$ と \mathbf{x}_t があるラベル $k \in K$ に属する尤度 $p(\mathbf{x}_t|k)$ を用いて求められる。また、この事後確率 $P(k|X_T)$ を用い、時間 T における最大事後確率 λ_T は式 (2) のように表わされる。

$$\lambda_T = \max_k P(k|X_T) \quad (2)$$

この最大事後確率は、判別正答率と等しくなることが期待され、データの信頼度と捉えることができる。RB-ARQ では、このデータの信頼度の閾値を λ として設定し、ユーザーは $\lambda_T > \lambda$ となるまで同じ思考想起を続ける。

3. 提案手法

前述の通り、従来、RB-ARQ における事前確率は、すべての文字で等しく設定されていた。それに対し本稿では、一般文章における文字出現確率と文字間遷移確率を、この事前確率に反映させる方法を提案する。文字出現確率とは、文章中でどの文字が多く出現するかの確率、文字間遷移確率は、全文字に対して、ある文字の次にどの文字が出現しやすいかを表した確率であり、膨大な文章コーパスにおける文字 N-gram の出現頻度を用いて作成することができる。文字 N-gram とは、文章を N 文字ずつの単位に区切り、各文字列の出現頻度を算出したもの [9] で、これを用いて事前確率は以下の式で定義される。

$$P(X_i) = \frac{N(X_i)}{\sum_{l \in K} N(X_l)} \quad (n = 1) \quad (3)$$

$$P(X_i|X_{i-n+1}^{i-1}) = \frac{N(X_{i-n+1}^i)}{N(X_{i-n+1}^{i-1})} \quad (n \geq 2) \quad (4)$$

X_i^j は、文字列 $X_1 X_2, \dots, X_M$ 中の i 番目から j 番目までの部分文字列を表し、 $P(X_i|X_{i-n+1}^{i-1})$ は、 $\{i - (n - 1)\}$ 番目から $(i - 1)$ 番目までの文字列が与えられたとき、 i 番目の文字が X_i となる条件付き確率、 $N(X_i^j)$ は i 番目から j 番目の部分文字列の出現頻度を表す。この文字 N-gram から作成した事前確率は、 $n=1$ のとき、各文字の単純な出現頻度確率となり（Uni-gram と呼ぶ）、 $n=2, 3$ のときは文字間遷移確率を表す（これらをそれぞれ Bi-gram, Tri-gram と呼ぶ）。本稿ではこの 3 つの確率をそれぞれ用いる。これにより、前の入力文字に応じて、次の入力文字に対する事前確率の変更を行い、入力確率が高いと思われる文字に高い事前確率を与えることで、それらの文字に対する事後確率が閾値 λ に達するまでの判別時間が短くなり、文章の入力性能が向上すると期待される。

4. 実験

4.1 実験データと前処理

本実験では、あらかじめ計測してある、被験者 3 名が P300 speller を用いて文字入力を行った際の EEG データを用いた。このデータはサンプリング周波数 1000Hz、刺激呈示間隔 175ms で、国際 10-20 法に基づく図 1 中の 5 電極 (Fz, Cz, Pz, O1, O2)、及び基準電極として左耳 A1、右耳 A2 に電極をそれぞれ貼付し計測されたものである。このデータを 20Hz にダウンサンプリングし、刺激（行または列）点灯後 0s から 0.65s まで 0.05s ごとに合計 14 データポイントをとり、 14×5 （電極）= 70 変数を特徴として抽出した。その後、線形判別分析（Linear Discriminant Analysis: LDA）を識別器として用い、学習には、20 文字分の入力データを用いた。呈示インターフェースには、図 2 のように文字が 7 行 10 列型に配置された日本語入力インターフェースを用いた。

ここで、インターフェース上の <小> は小文字入力時に使用する文字であり、例えば「つ」を入力する場合は、<小><つ> の順で入力を行う。また <BS> は、訂正機能（Back Space）を表し、この文字が入力された場合は直前の文字を消去する。文字 N-gram による事前確率は、‘日本語ウェブコーパス 2010 [10]’ を基に、インターフェースに合わせコーパス上のすべての文字をひらがなに直し、式 (4) を用いて算出した。なお、本インターフェースを使用する上で、文字と分けて入力される濁点・半濁点は、文章中でそれぞれ濁点・半濁点のつく文字の総数を用いて確率計算を行った。また <小> の確率算出には、同様に小文字の出現総数を用いた。

4.2 実験方法と評価指標

前述のあらかじめ計測された脳波データには‘P300’ と ‘non-P300’ のラベルが付いているため、テスト時にはそこからターゲット行/列の点灯時には P300 を、それ以外の時には non-P300 をランダムに取り出し、シミュレーションを行った。テスト文には Web ブログ、エッセイ、小説から抜粋した 200 文字前後の 3 つの長文を用い、各文の入力を 100 回繰り返した。本実験

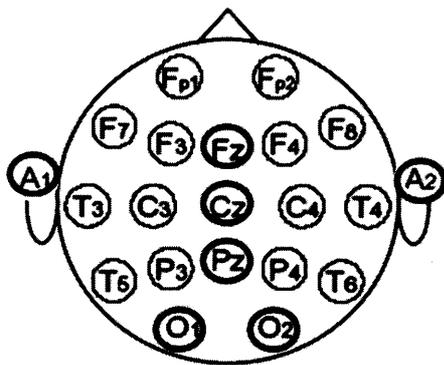


図1 使用電極

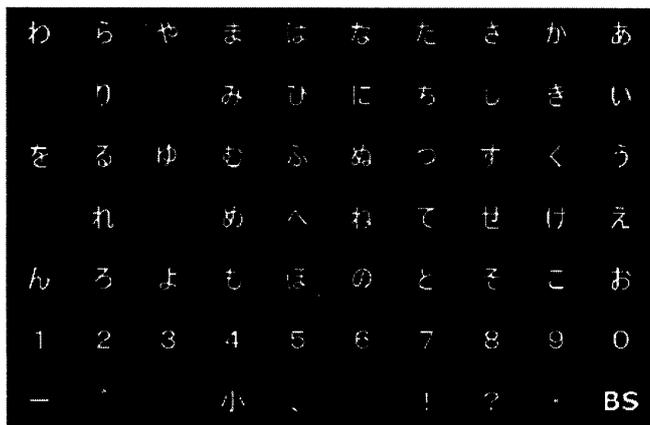


図2 P300 speller インターフェース

では、誤った文字が入力された場合、<BS> を用いて訂正を行い、正しい文を入力するものとし、RB-ARQ における閾値入を 0.9, 0.95, 0.99 の 3 種類に変化させて、以下の 2 つの方法で実験を行った。

実験 1: 事前確率を、Uni-gram, Bi-gram, Tri-gram それぞれを用いて設定し、従来法(事前確率を等確率に設定したもの、Equal と表す) との比較を行う。

実験 2: 事前確率を、Uni-gram, Bi-gram, Tri-gram それぞれを用いて設定し、<BS> と <BS> 以外の文字の入力が連続して 2 回以上繰り返された場合のみ、次文字の入力時に事前確率を等確率に設定する。

3 文に対する判別正答率と、単位文字あたりの点灯刺激数・入力所要時間の平均値を、それぞれ以下のように定義し、算出した。

$$\text{判別正答率} = \frac{\text{正解した文字数}}{\text{入力文字数}} \quad (5)$$

$$\text{単位文字あたりの点灯刺激数} = \frac{\text{全点灯刺激数}}{\text{入力文字数}} \quad (6)$$

$$\text{入力所要時間} = \text{単位文字あたりの点灯刺激数} \times \text{刺激呈示間隔} \quad (7)$$

判別性能に対する総合的な評価指標として、BCI の分野で最もよく用いられる性能指標の一つである Utility [11] を用いた。

$$U = \frac{(2P - 1) \log_2(C - 1)}{d} \quad (8)$$

Utility は、誤判別文字を消去する機能である <BS> が文字候補の一つにあることを想定したときに、完璧な文字入力を行う場合の情報伝達速度を示している。ここで、 C は判別クラス数(選択可能な文字の種類を示す。本実験では $C = 70$)、 P は判別正答率、 d は一文字あたりの平均入力所要時間である。ただし P が 0.5 以下の場合には、 $U = 0$ となる。

5. 結果と考察

実験 1 における判別正答率、単位文字あたりの点灯刺激数と入力所要時間を表 1 に、Utility を図示したものを図 3 に、実験 2 における判別正答率、単位文字あたりの点灯刺激数・入力所要時間を表 2 に、Utility を図 4 にそれぞれ示す。図 3,4 において縦軸は Utility の値、横軸は RB-ARQ の閾値を表している。

表 1 実験 1

	閾値	Equal	Uni-gram	Bi-gram	Tri-gram
判別正答率	0.9	0.810	0.794	0.818	0.65
	0.95	0.869	0.87	0.891	0.819
	0.99	0.918	0.933	0.95	0.939
点灯刺激数 [回]	0.9	78.9	68.0	55.2	41.9
	0.95	90.6	80.0	66.2	57.7
	0.99	110.8	100.6	85.4	83.0
入力所要時間 [秒]	0.9	13.8	11.9	9.7	7.3
	0.95	15.9	14.0	11.6	10.1
	0.99	19.4	17.6	14.9	14.5

表 2 実験 2

	閾値	Equal	Uni-gram	Bi-gram	Tri-gram
判別正答率	0.9	0.810	0.797	0.805	0.777
	0.95	0.869	0.869	0.88	0.859
	0.99	0.918	0.933	0.944	0.941
点灯刺激数 [回]	0.9	78.9	64.7	58.8	54.1
	0.95	90.6	76.5	69.8	63.6
	0.99	110.8	98.9	91.0	83.2
入力所要時間 [秒]	0.9	13.8	11.3	10.3	9.5
	0.95	15.9	13.4	12.2	11.1
	0.99	19.4	17.3	15.9	14.6

表 1 より、Uni-gram や Bi-gram の判別正答率は従来法(Equal) と同程度であるが、点灯刺激数が減少していることがわかる。しかし、閾値 0.9 の Tri-gram においては、点灯刺激数は減少しているものの、同時に判別正答率も大きく低下してしまっている。Tri-gram では、他手法に比べて特に、入力可能文字の事前確率の差が大きいため、事前確率の高い文字がターゲットとされた場合は大きな入力時間の短縮が見られる一方、設定された閾値が低くかつ事前確率の低い文字がターゲット

トとされた場合は、ターゲット文字の事後確率が上がる前に事前確率の高い文字が先に閾値を越えてしまい、誤入力となってしまったと考えられる。その結果訂正を行う<BS>の入力と誤入力が繰り返されることで、正答率が低下したと思われる。閾値が上がることでこの現象が起きる頻度が下がるため、表1において閾値0.95と0.99のTri-gramの正答率は向上していることがわかる。このことは、図3におけるUtilityでも確認できる。

一方、表2より、実験2のような工夫を行うことで、低い閾値でもTri-gramの正答率が向上していることがわかる。今回用いたテスト文は、長文であり、文字間のつながりが弱まると考えられる助詞も多く含まれていた。そのため、事前確率の低い文字がターゲットとして指定されることが多く、実験1においては、それらの部分で<BS>と誤入力が繰り返され、Tri-gramの性能が低下し、実験2ではそれらが改善されたと考えられる。また従来法とTri-gramを比較すると、1文字あたりの入力所要時間が4~5秒短縮されていることがわかる。それにより図4において、閾値0.9におけるTri-gramの性能が改善され、すべての閾値においてBi-gramよりUtilityの値が向上していることが確認できる。このUtilityの値に対して、従来法と提案する各手法(Uni-gram, Bi-gram, Tri-gram)間で、多重比較を考慮した(ボンフェローニの検定法)対応のあるt検定を行ったところ、統計的有意差($p=0.0104, 0.0021, 0.0042 < \alpha = 0.05/3$)が確認された。これにより、従来のように事前確率を等しく設定する場合と比較して、文字の出現確率を考慮することで性能は改善し、文字間のつながりやすさによる遷移確率を考慮することでさらに性能向上が見られることを確認した。

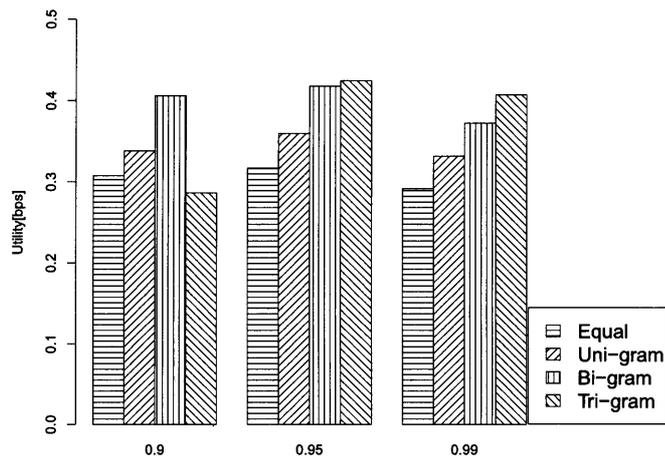


図3 実験 1

6. おわりに

本稿では、日本語文章におけるひらがな文字間の遷移確率を、RB-ARQの事前確率として組み込む手法を提案した。日本語

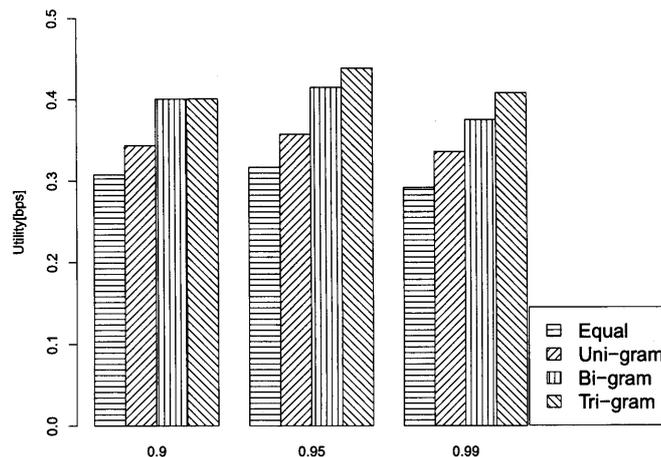


図4 実験 2

インターフェースによるP300 spellerの入力実験を行い、従来手法と比較して、正答率を保ちつつ、入力時間を短縮し、総合的な判別性能の向上が可能となることを確認した。今後は、オンライン実験や、<BS>情報を用いたさらなる性能向上に取り組んでいく予定である。

文 献

- [1] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, T. Vaughan, et al.: "Brain-computer interfaces for communication and control", *Clinical neurophysiology*, **113**, 6, pp. 767-791 (2002).
- [2] L. Farwell and E. Donchin: "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials", *Electroencephalography and clinical Neurophysiology*, **70**, 6, pp. 510-523 (1988).
- [3] U. Hoffmann, J. Vesin, T. Ebrahimi and K. Diserens: "An efficient p300-based brain-computer interface for disabled subjects", *Journal of Neuroscience methods*, **167**, 1, pp. 115-125 (2008).
- [4] R. Scherer, G. Muller, C. Neuper, B. Graimann and G. Pfurtscheller: "An asynchronously controlled eeg-based virtual keyboard: improvement of the spelling rate", *Biomedical Engineering, IEEE Transactions on*, **51**, 6, pp. 979-984 (2004).
- [5] H. Takahashi, T. Yoshikawa and T. Furuhashi: "A study on application of reliability based automatic repeat request to brain computer interfaces", *Proc. 15th Int. Conf. Neural Information Processing*, pp. 1013-1020 (2009).
- [6] Y. Kaneda, H. Takahashi, T. Yoshikawa and T. Furuhashi: "A study on application of reliability-based automatic repeat reQuest to p300 speller", *IEICE Technical Report*, **109**, 280, pp. 19-22 (2009).
- [7] K. Kukich: "Techniques for automatically correcting words in text", *ACM Computing Surveys (CSUR)*, **24**, 4, pp. 37-439 (1992).
- [8] J. Ullmann: "A binary n-gram technique for automatic correction of substitution, deletion, insertion and reversal errors in words", *The Computer Journal*, **20**, 2, pp. 141-147 (1977).
- [9] W. Cavnar, J. Trenkle, et al.: "N-gram-based text categorization", Vol. 48113, *Citeseer*, pp. 161-175 (1994).
- [10] <http://s-yata.jp/corpus/nwc2010/ngrams>.
- [11] B. Dal Seno, M. Matteucci and L. Mainardi: "The utility metric: a novel method to assess the overall performance of discrete brain-computer interfaces", *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, **18**, 1, pp. 20-28 (2010).