

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 14 日現在

機関番号：11501

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330224

研究課題名(和文) 符号理論に基づく視覚刺激系列を用いた脳波による高速文字入力システムの開発

研究課題名(英文) Development of high-speed EEG character input system by visual stimulus series based on code theory

研究代表者

深見 忠典 (FUKAMI, TADANORI)

山形大学・理工学研究科・准教授

研究者番号：70333987

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：本研究は、脳波を用いて高速にコンピュータに入力するための研究である。従来のP300 spellerでは、ユーザーに入力したい文字が点灯した際に点灯回数を数えるなどのタスクを与えておき、ディスプレイ上に格子状に配置された候補文字を行/列単位で点灯させ、タスク実行により出現するP300により文字特定を行う。我々は、点灯を行/列単位に拘らずに複数の文字を同時に点灯させ、なおかつ文字間の点灯/消灯パターンの相違を大きくすることで、高い文字識別精度で、識別までの所要時間を短縮する手法を着想した。結果として、従来法による文字入力結果に対して、我々の手法により約2.7倍の文字入力性能の向上が見られた。

研究成果の概要(英文)：This is the research to input a character into computers by EEG with high speed. In conventional P300 speller, on the condition that users are instructed to perform a task such as counting the number of lighting, candidate characters light up with one row or one column randomly and character identification is done by P300. We proposed the method to shorten a required time for identification with high accuracy. In this method, we conceived an ideal for the stimulus presentation which maximize differences of stimulus patterns assigned to each character without persisting a way of stimulus presentation based on unit of row/column. As a result, the number of input characters by using our method showed approximately 2.7 times larger than the conventional speller.

研究分野：生体情報処理，医用画像処理

キーワード：ブレインコンピュータインターフェイス 脳波 刺激パターン符号化 視覚刺激間隔

1. 研究開始当初の背景

脳波を用いた文字識別には、人間の認知や判断を反映する脳波である P300 が用いられる。これは、ディスプレイ上に複数の候補文字を配置し、標的文字を呈示（点灯）させた際、被験者にその文字を強く意識させることで、顕著に発生する。通常、自発脳波（生きている限り出現）に重畳して観測されるため、信号処理技術により P300 を検出することで、標的文字の識別に使用される。

文字呈示法として、Farwell と Donchin[1]により提案された P300 スペラが最も知られており、6×6 の文字に対して行もしくは列の単位でランダムに点灯する。よって、1 試行あたり 12 回（6 行+6 列）の呈示を必要とする。この呈示法は、現在でも広く用いられており、報告も多くなされている。また、世界で初めて製品化された g.tec 社 intendix[2]もこの方式を採用しているが、1 分間に 5~10 文字程度の入力に留まる。その原因として、標的文字が点灯する際、行または列で固定された複数文字と伴に点灯するため、標的文字と異なる文字を誤って検出するリスクが増加し、識別精度の低下に繋がるためであると考えられる。そこで、最近では 1 文字ずつ高速かつランダムに点灯させる方法が提案されている[3]。これにより識別精度は向上するものの、全候補文字を点灯させる（1 試行あたり 36 回）必要があるため、P300 スペラに比べ、より文字呈示に時間を要する。このように、高精度の文字識別と識別に至るまでの所要時間短縮の両立に課題が残されているというのが現状である。

2. 研究の目的

本申請課題による研究期間内において以下の 3 点を実施し、研究期間終了後には、一つの BCI システムとして稼働させることを目指す。

1. 脳波解析アルゴリズムの確立：短時間に多数の文字入力を行うためには、提案する文字呈示法を最大限に生かす標的文字推定のための解析アルゴリズムの確立は必須である。一般的には、識別精度向上のため、各個人における P300 特徴量（振幅値など）に対し、機械学習が用いられるが、脳波は人間の状態（疲労など）により時々刻々変化するため、常に最良の識別器が得られる保証がない上、学習用データ取得に時間を要するため、本研究では、機械学習を用いない方針である。

3. 研究の方法

(1) 最適な呈示パラメータの獲得

提案法の文字呈示において用いられる次の a, b の最適パラメータや文字呈示パターンを脳波計測実験による識別精度及び所要時

間の観点から獲得する。

① 最適な符号長及び候補文字間の符号間距離の決定（本研究では、消灯/点灯を 0/1 として符号化する）

② 最適な文字点灯率の決定：符号間距離最大化のため点灯頻度を高くする必要がある一方、P300 は低頻度刺激で顕著に出現することが知られており、計測実験により最適な点灯率を取得する。

(2) 刺激系列の生成

これまでの他の研究者の多くは、分類器や前処理部分に計算コストの高いアルゴリズムを使用し、文字判別精度の向上を目指してきた。その結果、判別に至るまでの時間が延長することで、脳波計測時間も増大し、被験者への負担も重くなっている。こうした計算コストの高いアルゴリズムの使用は、標的文字と非標的文字に対する応答波形が似ており、それに伴う特徴量が類似に起因しているためであると考えられる。よって、類似する特徴量から、標的・非標的を区別するために、必然的にアルゴリズムが複雑となる。

そこで本研究では、発想を転換し、文字の刺激呈示方法に着目することで、各文字の区別を容易にする脳波を出現させるための点滅パターン作成に着手した。

具体的には以下の 2 点を組み合わせた評価値を用いて、点滅パターンを作成した。

①各文字 1 セットにおける点灯数のうち、各文字の同時点灯回数を減らす

②文字の点灯間隔を広くする

P300 Speller では、1 試行（6 行+6 列）で 1 文字当たり 2 回の点灯に留まり、標的文字とそれ以外の文字との同時点灯回数が最大 1 回となる。この時、顕著な P300 が出現するため、標的文字でない文字を標的文字とする誤った識別をする可能性がある。

そこで、1 試行当たりの文字の点滅回数を増やし、かつ各文字の同時点灯回数を少なくすることで、標的文字と非標的文字との点灯消灯パターンの差を大きくすることができると考えられる。本研究では、次式に示すハミング距離を最大化するような点灯消灯パターンを決定する。

$$h(i, j) = \sum_{k=1}^S |i_k - j_k| \quad \text{式 1}$$

ここで $h(i, j)$ はハミング距離、 i, j は異なる文字、 S は 1 試行の点滅数を示している。次に一文字とその他全ての文字とのハミング距離を求め、式 2 に示す平均化させた平均ハミング距離 f_1 を定義した。

$$f_1 = \frac{1}{N-1} \sum_{l=1}^N h(c_k, c_l) \quad \text{式 2}$$

ここで、 N は文字数、 c は文字、 $h(c_k, c_l)$ は k 番目と l 番目の文字のハミング距離を示す。

しかしながら、各文字の同時点灯回数が少なくなっても1試行当たりの文字の点灯回数が増えてしまえば、標的文字の刺激呈示が高頻度となり、P300の振幅値の低下を招く。これにより、振幅値を用いた標的文字の識別が困難になると考えられることから、P300の振幅値が大きくなるように標的文字の刺激呈示を低頻度化させる必要がある。標的文字の刺激呈示を低頻度化させるために、点灯間距離の期待値を f_2 として、以下の式により算出する。

$$f_2 = \frac{1}{F-1} \sum_{k=1}^{F-1} F_{Nk} \quad \text{式 3}$$

ただし、 F は1セットの点灯数、 F_{Nk} は点灯間の距離（消灯数）を示している。

こうして得られた異なる2つの評価値を $f=f_1+\alpha f_2$ により統合する。 α は2種類の距離間の割合、すなわち重みであり次式により決定する。

$$\alpha = \frac{\bar{f}_2}{\bar{f}_1} \quad \text{式 4}$$

重みは、全文字に対して多数回ランダムに符号化し平均ハミング距離の平均値

\bar{f}_1 と点灯間距離の期待値の平均値 \bar{f}_2 の2つの平均値を求めるシミュレーション実験を行うことで算出した。これにより、2つの異なる距離を統合した次式の評価関数 d_i を最大化することにより符号割り当てを行った。

$$d_i = \sqrt{\alpha^2 f_{1i}^2 + f_{2i}^2} \quad \text{式 5}$$

ただし、 i は文字を示している。

(3) 標的文字の推定方法

バンドパスフィルタ（1～7 Hz）を適用した後、P300の推定を行う。本研究では、2通りの処理による波形を求めた。1つは、一般的に広く用いられている、刺激を複数回繰り返して呈示し、刺激呈示時刻を揃えて、点灯文字に対してのみ加算平均処理を施すものである。

$$y_a(t) = \frac{1}{T_F} \sum_{i=1}^N \sum_{j=1}^M kx_{ij}(t), \quad k = \begin{cases} 1 & \text{点灯時} \\ 0 & \text{消灯時} \end{cases} \quad \text{式 6}$$

N は試行数、 M は1試行における点滅数、 $x_{ij}(t)$ は i セット目の j 番目の点滅時におけるバンドパスフィルタ後の信号、 T_F は1試行内の1文字の点灯数である。もう一つは、文字点灯時の加算平均処理に加え、文字消灯時も加算平均処理に組み入れた方法、すなわち文字点灯時の応答から文字消灯時を減算した波形を用いる。算出は以下の式により行われる。

$$y_b(t) = \sum_{i=1}^N \sum_{j=1}^M x_{ij}(t) \left(\frac{k}{N \times T_F} - \frac{l}{N(M - T_F)} \right),$$

$$k = \begin{cases} 1 & \text{点灯時} \\ 0 & \text{消灯時} \end{cases}, \quad l = \begin{cases} 0 & \text{点灯時} \\ 1 & \text{消灯時} \end{cases}$$

式 7

N は試行数、 M は1試行における点滅数、 $x_{ij}(t)$ は i 試行目の j 番目の点滅時におけるバンドパスフィルタ後の信号、 T_F は1試行内の1文字の点灯数である。

(4) 標的文字の判別方法

本研究における識別方法は、各試行におけるP300潜時の200msから500ms間における最大振幅値を用い、最大振幅値を示した文字を標的文字とする。リアルタイムによる実行を実現するためには、標的文字を確定した時点で次の文字入力に移る必要がある。そのため、標的文字を確定する基準として、式8に示すようにここでは最大振幅値と2番目に大きい振幅値との差を用いる。

$$\frac{amp_1 - amp_2}{amp_{max} - amp_{min}} > D \quad \text{式 8}$$

ここで amp_i は、区間内で i 番目に大きい振幅値を示し、 D は文字判別条件を示している。式8は、最大振幅値と2番目に大きい振幅値の差だけでなく、最大振幅値と最小振幅値の差により正規化することで、各文字の振幅値のばらつきも考慮している。加算回数が少ない時点では、ノイズを多く含んでいることが多く、その変動により、誤って標的文字と判別するリスクを抑えることができる。

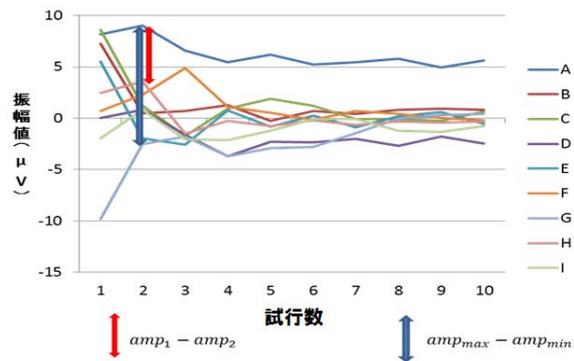


図1. 試行回数と振幅値変化の一例 (A～Iは候補文字を表している)

(5) 性能評価指標の定義

本研究では、入力文字推定における性能評価について導入した4つの評価指標について述べる。

まず推定した標的文字の正答率 A を式9に定義する。

$$A = \frac{1}{N_E} \sum_{i=1}^{N_E} target_i \quad \text{式 9}$$

ここで、 N_E は実験回数、 $target_i$ は文字が標的文字と識別された回数を示す。次に文字識別までの時間 T (s) を式 10 に定義する。

$$T = \frac{1}{N_E} \sum_{i=1}^{N_E} (M \times I \times S_i) \quad \text{式 10}$$

N_E は実験回数、 M は1試行における点滅回数、 I は刺激間時間 (0.6s)、 S_i は i 回目の実験における文字識別までに要したセット数を表している。

第3の指標として、標的文字を1分当たり入力することができる文字数 C は式 11 で定義する。

$$C = \frac{60}{T} \times A \quad \text{式 11}$$

最後に式 9 の文字正答率式 A と式 10 の識別時間 T を用いて、ビットレート B (bit/min) を式 12 により計算する。なお、 N_C は実験で用いる文字数である。

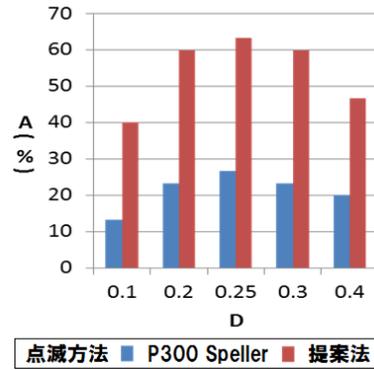
$$B = (\log_2 N_C + A \log_2 A + (1-A) \log_2 \frac{1-A}{N_C-1}) \times \frac{60}{T} \quad \text{式 12}$$

4. 研究結果

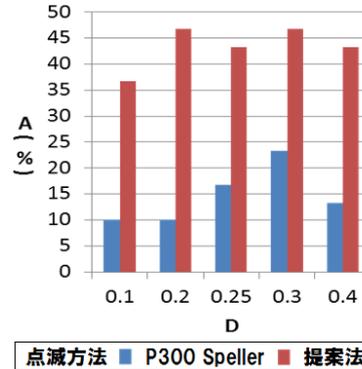
本研究における文字数は、符号間距離との関係を詳細に観察するため、一般的な P300 Speller で用いられている文字数を減らした 3×3 の 9 文字において実験を行う。また、文字呈示間隔についても多くの研究で用いられているものよりも長い 0.6s とした。

結果は、被験者男性 3 名 (年齢: 24.7±0.58 歳) の平均値として示した。また、式 8 中の D は、0.1, 0.2, 0.25, 0.3, 0.4 から最適値を求めた。標的文字の正答率を図 2 に示す。P300 Speller の点滅法に比べ、提案法では正答率が向上した。 $y_a(t)$ を用いた場合は、 D の値が 0.25 のとき最大となり、提案法は 63.3%、P300 Speller は 26.7%であった。また、 $y_b(t)$ を用いた場合は、 D が 0.3 のとき最大となり、提案法は 46.7%、P300 Speller は 23.3%であった。

文字識別までの時間について図 3 に示す。 $y_a(t)$ を用いた場合、 D が小さい時、提案法は P300 Speller に比べ時間を要した。 D が大きくなるにつれ、提案法の識別時間の方が短くなった。 $y_b(t)$ を用いた場合は、 D の値に関わらず提案法の方が短い結果を示した。

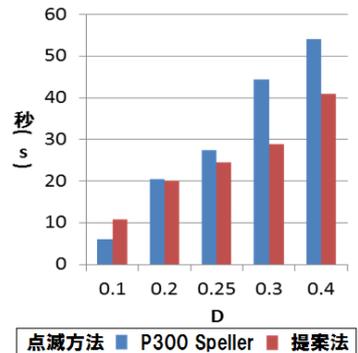


(a) $y_a(t)$ による結果

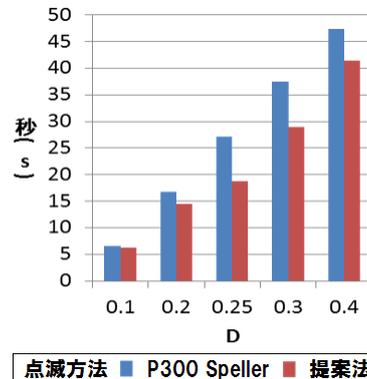


(b) $y_b(t)$ による結果

図 2. 標的文字の正答率



(a) $y_a(t)$ による結果



(b) $y_b(t)$ による結果

図 3. 文字識別までの時間

図 4 には、1 分あたりの入力文字数を示している。提案法の方が、大きい結果を示した。また、 D が大きくなるにつれて、文字

数が減少する傾向にあった。 $y_a(t)$ を用いた場合、 D が0.1のとき最大となり、入力文字数は提案法で2.3文字/分、P300 Spellerで1.6文字/分であり、提案法の方が約1.4倍多く入力できる結果となった。また、 $y_b(t)$ を用いた場合では、 D が0.1のとき最大となり、提案法で2.4文字/分、P300 Spellerで0.9文字/分であり、提案法の方が約2.7倍多くの文字を入力することができる結果となった。

最後に、ビットレートの結果を図5に示す。提案した点滅方法の方が、高いビットレートを示した。 $y_a(t)$ 、 $y_b(t)$ を用いたいずれの場合も D が0.2で最大となり、ビットレートはそれぞれ3.4bit/分、2.5bit/分であった。

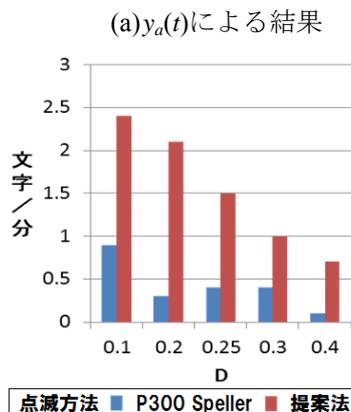
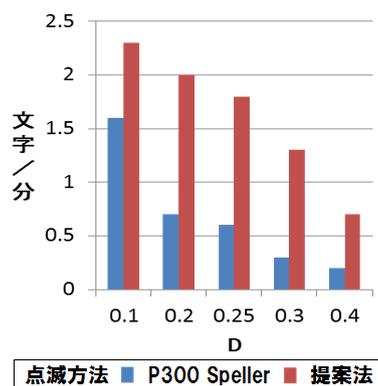


図4. 1分あたりの入力文字数

考察

正答率においては、 D が小さいもしくは大きすぎると正答率が悪くなる傾向にあった。この理由は、 D が小さいと非標的文字を誤って識別してしまう可能性が高くなるのが原因だと考えられる。また、 D が大きすぎると本研究で設定した最大試行数では、最大振幅値と2番目に大きい振幅値との差が、 D 以上の値になりにくくなるのが原因であると考えられる。そこで、今後適切

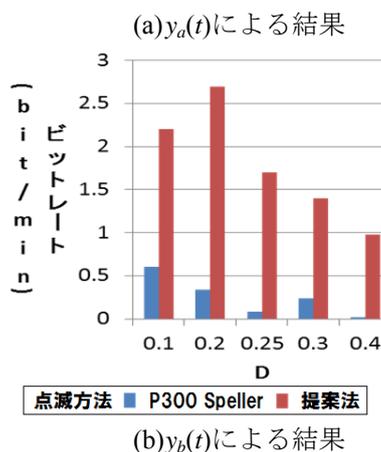
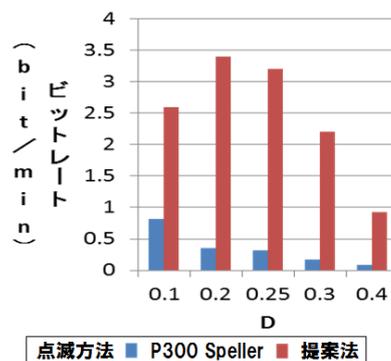


図5. ビットレート

な D の値を設定することが必要となる。次に、提案法を用いた正答率が高かった理由に、P300 Spellerに比べ、標的文字の振幅値が早期に最大になったこと、標的文字と2番目に高い振幅値との差が大きくなった点が挙げられる。

次に、文字判別までの時間については、提案法の方が、P300 Spellerに比べ、早期に高い振幅値を示していた。本研究では、判別条件に文字の振幅値を利用しているため、識別までの時間短縮を図るためには、文字同士の振幅値の関係が非常に重要になる。このため、提案法の方が、P300 Spellerに比べ、標的文字と非標的文字との振幅値差を大きくすることが可能となり、識別までの時間が早まったと考えられる。図3(a)において、本研究で用いた判別条件 D の値が0.1のとき、提案法の方がP300 Spellerに比べ、判別までの時間が遅くなっていた。これは、1試行終了後に標的文字の判定を行っているため、P300 Spellerによる点滅法は、1試行あたり6回の点滅であるのに対し、提案法による点滅方法は、12回の点滅となっていることが原因であると考えられる。

1分あたりの入力文字については、提案法の方がP300 Spellerに比べ、入力可能文字数が大きくなった。この理由は、本手法がP300 Spellerに比べ、標的文字の正答率が高く、識別までの時間が全体的に短いことが挙げられる。

最後にビットレートに対して考察を行う。

$y_b(t)$ が $y_a(t)$ に比べ、ビットレートが減少した理由は、正答率が大きく影響していると考えられる。 $y_b(t)$ は、判別時間は短いものの正答率が悪く、ビットレートが悪化している。よって、ビットレートは判別時間よりも正答率に大きく依存するものと考えられる。

ビットレートが低いのは、実験での文字の点滅間隔と1試行毎の文字判定に原因があると考えられる。すなわち本研究では、P300の応答を波形レベルで確認するため、刺激時間を600msとしており、BCI研究で多く用いられている175msに比べ長い。これはビットレートを低下させる一因である。また、一試行12回の刺激に対する応答から標的文字を推定しているが、試行途中であっても標的文字の判定が可能であると判断した場合は、そこで打ち切り次の文字に移行することでビットレートは向上する。

今回の実験結果と他の研究者による結果を比較するためには、実験条件を一致させる必要があるため、一概に比較はできない。

結論

本研究では、脳波を用いた高速文字入力法として、P300 Spellerのような文字を1行または1列で非効率に呈示する方法を改善し、新たな文字刺激呈示法を考案した。すなわち、文字の点灯/消灯を1/0で符号化した後、各候補文字に割り当てた符号間距離と同一文字の刺激呈示間隔の両方を大きくする刺激系列を算出し、それをもとに呈示された刺激により得られた脳波を解析、文字入力に関する指標により評価した。

提案法では、正答率、文字識別までの時間、1分当たりの入力文字数とビットレートにおいてP300 Spellerに比べ良い結果を示した。1分当たりに入力することができる文字数は、標的刺激応答の加算平均処理において、P300 Spellerでは1.6文字/分に対し、提案法では2.3文字/分と約1.4倍であった。また、全応答を用いた平均処理において、P300 Spellerでは0.9文字/分、提案法は2.4文字/分となり、約2.7倍の性能向上が見られた。

<引用文献>

[1] L.A. Farwell, E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials.", *Electroencephalogr. Clin. Neurophysiol.*, vol.70, no.6, pp.510-523, 1988.

[2] <http://www.intendix.com/>

[3] C. N. Gupta, R. Palaniappan, R. Paramesran, *J. Cognitive Biometrics*, vol. 1, no. 1, pp.26-38, 2012.

5. 主な発表論文等

[雑誌論文] (計 3 件)

① T.Fukami, J.Watanabe, F.Ishikawa, "Robust estimation of event-related potentials via particle filter", *Computer Methods and Programs in Biomedicine*, (査読有), vol.125, pp.26-36, 2016.

② T.Fukami, K.Ishihara, F.Ishikawa, "Preliminary Study for Extraction of P300 and SSVEP by Stimulus Presentation Using Phase Inversion Technique in Hybrid BCI.", (査読有), *Proc. of 37th Annual International Conference of IEEE Engineering in Medicine and Biology Society (EMBC2015)*, pp.63-66, 2015.

③ J.Toma, T.Fukami, T.Shimada, "Character Identification by Maximizing the Difference between Target and Non-target Responses in EEG without Sophisticated Classifiers", (査読有), *Proc. of 35th Annual International Conference of IEEE Engineering in Medicine and Biology Society (EMBC'13)*, pp.2243-2246, 2013.7.

[学会発表] (計 4 件)

① 深見忠典, 渡辺 潤, 石川 史人, "粒子フィルタを用いたP300脳波成分の推定", 計測自動制御学会システム・情報部門学術講演会2015, p.1330, 2015.11.20. (於 北海道, 函館市, 函館アリーナ)

② 渡辺 潤, 澤木 綾太, 深見忠典, 島田 尊正, 石川 史人, "リアルタイムBCIシステム構築を目的とした粒子フィルタによるP300成分の抽出", パーティクルフィルタ研究会, (4), 2015.5.23. (於 山形県, 米沢市, 山形大学工学部)

③ 澤木 綾太, 渡辺 潤, 遠間 洵也, 深見忠典, 石川 史人, "視覚刺激系列の最適化に基づく脳波による文字入力の時間短縮", 第54回日本生体医工学学会大会, P3-2-12-B, 2015.5.9. (於 愛知県, 名古屋市, 名古屋国際会議場)

④ 渡辺 潤, 遠間 洵也, 深見忠典, 島田 尊正, 石川 史人, "粒子フィルタを用いた事象関連電位からのP300成分抽出", 第53回日本生体医工学学会大会, O2-03-3, 2014.6.25. (於 宮城県, 仙台市, 仙台国際センター)

6. 研究組織

(1)研究代表者

深見 忠典 (FUKAMI, Tadanori)

山形大学・大学院理工学研究科・准教授
研究者番号： 70333987