

機関番号：17201

研究種目：若手研究 (B)

研究期間：2009～2010

課題番号：21700583

研究課題名 (和文) 表面筋電信号による代用発声システムに関する研究

研究課題名 (英文) Study of Substitute Vocalization System based on the Surface Electromyogram

研究代表者

福本 尚生 (FUKUMOTO HISAO)

佐賀大学・工学系研究科・助教

研究者番号：60346872

研究成果の概要 (和文)：本研究の目的は、音声によってコミュニケーションをとることができない場合に、声を発することなく、口唇周辺の表面筋電信号から発話を認識することにある。しかし、無発声動作に慣れていない人では認識率が低下する。そこで、複数の被験者から得られた筋活動値のパターンを、自己組織化マップ(SOM)のアルゴリズムを使用してクラス分けし、母音に対する標準的な筋活動値のパターンを決定し、訓練するシステムを提案した。

研究成果の概要 (英文)：The purpose of this work is to achieve unvoiced speech recognition using Surface Electromyogram (SEMG) of the facial muscles when voiced communication is impossible. However, the recognition accuracy was low for some speakers who were beginner of unvoiced speech. To address this issue, we have examined the classification of the activity values measured from multiple speakers using Self-Organizing Map (SOM) algorithm and determined the standard pattern of the activity values for each vowel. In addition, the author has proposed a training system for unvoiced speech.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009年度	2,000,000	600,000	2,600,000
2010年度	900,000	270,000	1,170,000
年度			
年度			
年度			
総計	2,900,000	870,000	3,770,000

研究分野：総合領域

科研費の分科・細目：人間医工学 リハビリテーション科学・福祉工学

キーワード：ヒューマンインターフェース、表面筋電信号、代用発声、パターン認識

1. 研究開始当初の背景

音声は我々にとって非常に重要な情報伝達手段であるが、喉頭癌などで声帯を切除された人々は、音声を発声することが困難であり、音声によって意思を伝えることが不可能である。このような障害を持つ方は高齢者に多いとも言われており、高齢社会を迎える中、こうした障害を持つ人が増えてゆくと思われる。また、疾患により体力が著しく衰えた場合にも、発声が困難となる場合もある。

声を出すという動作には想像以上の体力を要する。さらに、静寂な環境が要求される場合においては、声を出して会話することが制限される場合もある。

発声自体が困難な場合の代用発声法は従来多くの方法が検討されており、その代表的なものに食道発声法や電気式人工喉頭などがある。しかし、食道発声法には独特な技術の習得が必要であり、既存の電気式人工喉頭は、機器を手で把持し下顎の下部に押し当て

て使用するため、片手が拘束されてしまうという問題がある。

これに対して、比較的自然的な方法で、特殊な訓練を要しない代用発声法として、口唇や舌の動きを支配している筋肉の動きを検出し、発話を認識する方法が検討されている。口唇周辺の表面筋電信号を用いた代用発声システムの研究としては、既に多くの研究報告がなされているが、いずれの報告にも母音に関する認識結果が中心であり、子音を含む音節の認識は未だに困難である。

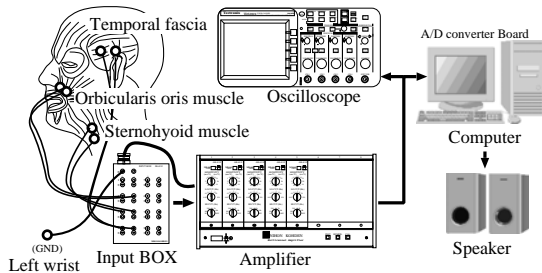


図1 無発声母音認識システム構成

2. 研究の目的

音声によりコミュニケーションをとることができない場合に、声を発することなく、口唇周辺の表面筋電信号から発話を認識することを最終目標としている。図1にシステムの構成を示す。口唇周辺の複数の筋肉に貼付した電極から筋電信号を取得し、増幅器により増幅、デジタル化した後、コンピュータによって筋電信号を基にした発話認識を行い、必要に応じて代用発声するというシステムである。本研究の目的は、母音の発話における筋動作の標準的なパターンを見だし、バイオフィードバックを積極的に利用した訓練システムを開発することである。また、安定的に無発声動作が可能になった被験者において、子音を含めた認識へ向けた検討を行うことである。

3. 研究の方法

(1) 計測部位

①母音認識における計測部位…母音の認識においては、先行研究などを参考に計測部位を検討した結果、本研究では、口輪筋(CH1)、側頭筋膜(CH2)、胸骨舌骨筋(CH3)の三カ所の周辺で筋電信号を計測した。

②子音も含めた認識における計測部位…子音も含めた認識を検討する際には、母音認識における計測部位に加え、新たに、茎突舌骨筋(CH4)の周辺で筋電信号を計測した。

(2) 計測方法

表面筋電信号を計測するための電極には、皿電極(NT-211U, 日本光電)を使用し、

双極導出とし、アース電極は手首や肘に貼付した。電極で検出された筋電信号は生体アンプ(AB-611J, 日本光電)によって1000倍に増幅され、15Hz~300Hzのバンドパスフィルタを通し、AD変換ボード(PCI-3166, インターフェース)を用い、サンプリング周波数1kHz、12bitのデジタル値としてコンピュータ(Dimension2400, DELL)に記録する。

記録された筋電信号は発話認識処理を行い、認識結果を文字として画面に表示するとともに、フリーソフトの音声合成ソフト(easy speech)を用いて音声として出力する。

(3) 信号処理方法

各部位の筋電信号の取得については、あらかじめ発話していない平常状態の筋電位の平均値を取得しておき、平均値の6倍を超えた場合に発話が始まったと判定する。発話が始まったと判定された後、500ms間の筋電位を取得し、50ms毎にRMS値を求める。RMS値を求める場合の時間間隔については、先行研究などを参考に50msに設定した。得られたRMS値の最大値をその発話における筋活動値とみなし、三カ所の部位における筋活動値パターンを用いて発話認識を行う。

これらの信号処理の流れを図2に示す。ここでは、/a/と発話した例を示しているが、/a/においては胸骨舌骨筋周辺(CH3)の筋が大きく活動する。また、筋活動の個人差を吸収するため、計測開始前にあらかじめ発話動作を複数回行い、各部位での筋電信号の最大値の平均値を求め、その値で正規化を行った後にRMS値を求めている。

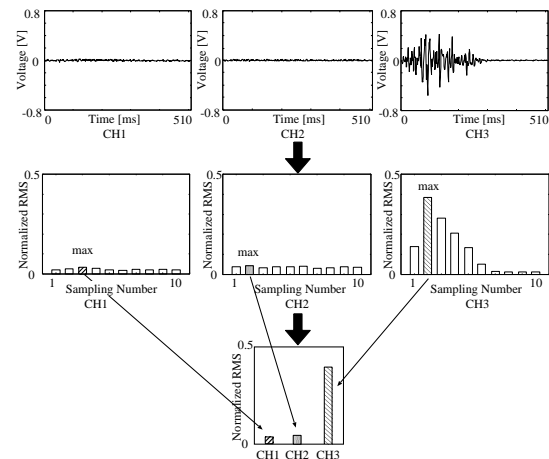


図2 筋電信号から筋活動値への信号処理

(4) 母音認識方法

母音の認識においては、前述した処理で得られたRMS値の最大値である筋活動値に対して、閾値を設定し閾値を越えた場合にその筋が活動したと判定する。したがって、筋活動の有無の二値化パターンによって母音の認識を行なうが、その閾値の決定に関しては複数の被験者で実験的に検討し0.125とした。

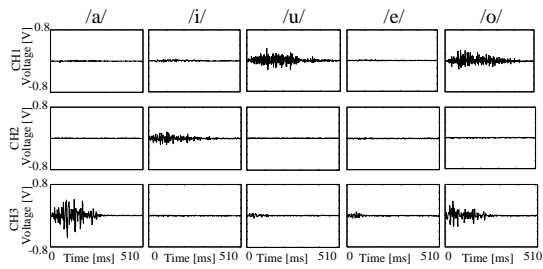


図3 母音の筋電信号の例

図3に各母音の発話における各部位の筋電信号の例を示す。このように各発話で筋の活動パターンが異なるので、筋活動値に閾値を設けて二値化パターンとして認識を行うことができる。表1に各発話と二値化パターンを示す。

表1 母音の二値化パターン

vowel	CH1	CH2	CH3
/a/	0	0	1
/i/	x	1	0
/u/	1	0	0
/e/	0	0	0
/o/	1	0	1

表中では、筋が活動した場合を「1」と表し、筋が活動していない場合を「0」と表している。/i/に関しては、人によって口輪筋周辺(CH1)の筋肉が動く場合と、動かない場合があったので、/i/の認識にはCH1を使用しないこととし、「x」と表している。

(5) 自己組織化マップ

自己組織化マップ(Self-Organizing Maps: SOM)は、T. Kohonenが提案したアルゴリズムであり、ニューラルネットワークアルゴリズムの一つである。図4にSOMの学習によるマップ形成の模式図を示す。アルゴリズムの詳細は文献に譲るが、SOMは多次元の入力を、二次元のニューロン配列上へマッピングを行う。入力ベクトルは、マップ上の全てのニューロンと重みベクトルによって結合されている。入力ベクトルと最もユークリッド距離の近い重みベクトルを持つニューロンが活性化(選択)され、学習過程では、その最も距離の近いニューロンの重みを、入力ベクトルとの誤差に基づき学習を行う。ここで、学習が行われる近傍範囲を決定する近傍関数によって、選択されたニューロンの近傍に関しても、同様の学習を行う。この近傍範囲と学習に用いる学習率を時間と共に小さくすることで、マップが形成される。

多数の入力ベクトルに対して、十分な繰り返し回数で上記の学習を行うことで、多次元の入力ベクトルをその類似性に応じて、二

元の空間に写像することができる。したがって、二次元のマップ上で近傍に設置されているニューロンが持つ重みベクトルは互いに類似しており、離れたもの同士では大きく異なるように学習される。

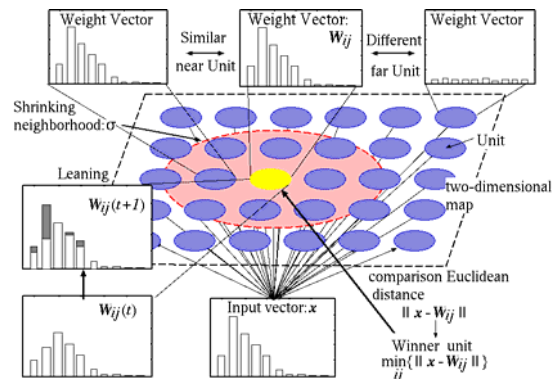


図4 SOMのマップ形成の模式図

(6) 子音も含めた認識

一部の子音も含めた発話認識では、筋動作にわずかな違いが見られることが予備実験などですでに明らかになっている。しかし、わずかな筋動作の違いを抽出するためには、十分に安定した筋動作が必要である。筋動作が不安定な場合には、後述する訓練システムなどを利用して、十分に無発声動作に慣れる必要がある。子音が含まれた発話では、唇の動きや舌の動きに母音とは異なる特徴が出る。例えば、/ta/や/na/と発話する場合には、/a/と比較すると舌の動きに特徴がでる。このことが、胸骨舌骨筋周辺(CH3)の筋活動の時間遅れとして計測される。また/ma/や/wa/の発話では、/a/の場合に比べて唇を一度閉じる動きが加わるために、口輪筋周辺(CH1)の筋活動がわずかに大きくなる。これらの微小な変化を判別することによって認識を可能にする。ここでは、閾値を設定して認識を行った。

(7) 訓練システム

無発声動作に慣れていない使用者では、筋活動値自体が低い場合や、安定しないなどの理由で母音の認識率は低い傾向にある。しかし、システムを使用している間に、筋活動値が高くなり、筋動作の個人差による多少の違いも、バイオフィードバックにより修正され、認識率が向上するケースが多く見られた。そこで、バイオフィードバックを積極的に利用するために、訓練システムを構築する。この訓練システムには、SOMで学習したマップを組み込み、使用者は、自分の発話動作の筋活動パターンと、標準的な筋活動パターンとの違いを、二次元のマップ上での位置関係から推測することができる。同時に、筋活動パターンを棒グラフとしても確認できるようにし、使用者は、標準的な筋活動値に近づくように

発話動作を訓練する。

4. 研究成果

(1) 無発声母音認識における検討

母音を認識する為には、口輪筋(CH1)、側頭筋膜(CH2)、胸骨舌骨筋(CH3)の三つの筋電計測で認識が可能であるが、発話の際に使用される口唇周辺の筋肉の筋電信号が、個人によってバラつくことが分かっており、このバラつきが大きいものなのか、小さいものなのかを本システムを初めて使用する被験者6名のデータで検討した。各被験者に母音をそれぞれ30回ずつ発話してもらい、この時の筋活動値の平均とその分散を図5に示す。

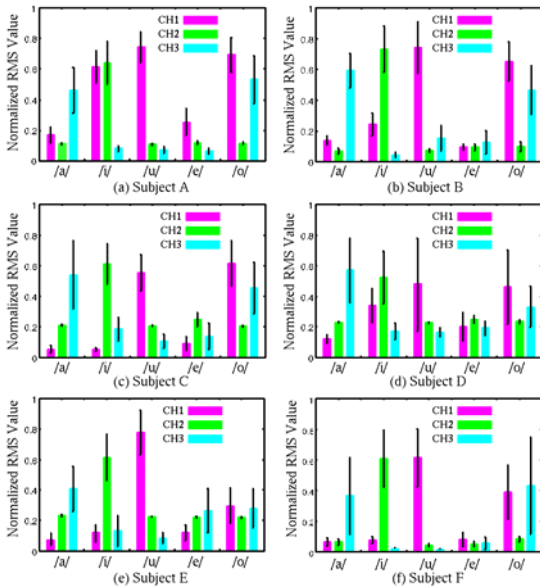


図5 被験者6名の筋活動値パターン

図5から、/i/の発話におけるCH1を除いては、筋活動値パターンは似た傾向を示しているが、無発声動作が不安定で、分散が大きく広がっている被験者もいることがわかる。このバラつきを含むデータを評価する際に、ニューラルネットワークの一種であるSOMを用いて解析を行った。バックプロパゲーションによる機械学習による認識では、無発声時の口唇動作が安定しない為に学習が十分に収束せず、認識率が低い結果が得られた。SOMでの解析では、安定しない動作においても、被験者間で同じような筋活動パターンが見られることが明らかになり、ある程度標準的な動作を抽出できた。

さらに、検討を重ねた結果、システムにある程度慣れた別の被験者6名で、合計4,050個の母音計測データを用いてSOMによって作成したマップを図6に示す。

このマップにおいては、トラスマップを採用しており、マップの上下、左右はそれぞれ連続している。ここで、マップ上で色分けされている部分が、各位置での重みのパターンを閾値によって母音の判定をした結果を

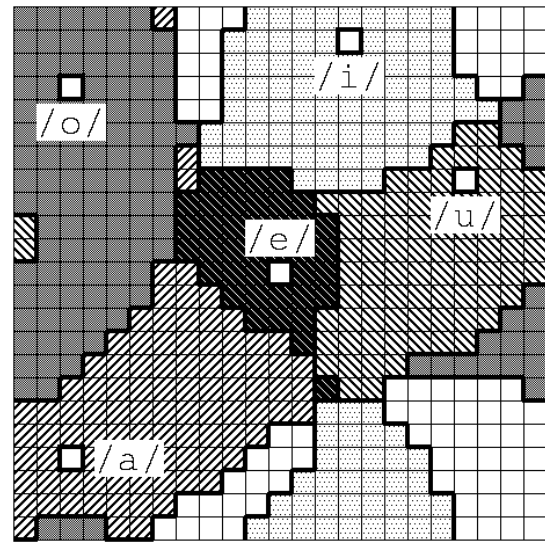


図6 SOMによって作成されたマップ

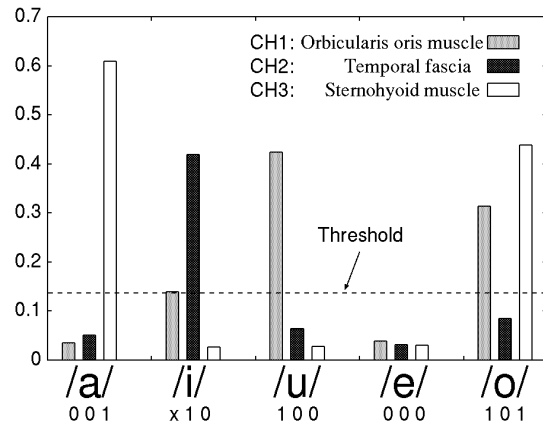


図7 抽出された各発話の標準パターン

示している。マップ上に各発話がグループ化されていることが分かる。また、このマップに学習に用いたデータを逆写像した結果、各発話において最も多くの計測データと一致する重みを持つユニットの重みベクトルを標準パターンとして決定し、マップ上の位置を図6の□で示すとともに、その標準パターンを図7に示す。このようにして得られた標準的な動作を基に、使用者の口唇動作の訓練によるバイオフィードバックを積極的に利用した方法を検討した。

(2) 無発声母音発話訓練システム

これまでに複数の被験者で母音の認識実験を行ったが、筋電信号のバラつきのために認識結果も大きく変化する結果を得た。これは、声を出さずに発声することに対する慣れや口唇動作の個人差などが原因と思われ、初めてシステムを利用する被験者の認識率はかなり低い結果であった。しかし、試行を繰り返す毎に認識率は高くなっていった。つまり、人間がシステムの特性に応じて認識しやすい動作を学習していくバイオフィードバックがなされたものと考えている。そこで、複数の被験者の筋電計測データから、平均的な口唇の動きである標準パターンを決定し、この標準パターンに近づくように訓練する

システムを構築した。そのシステムの外観を図8に示す。図の右側にはSOMのマップを示しており、被験者は、自分の発話動作パターン（白い点で表示）と、標準パターン（色別に表示している）の違いを二次元マップとして視覚的に確認できる。また、被験者は標準パターンと自分の発話パターンを左下の棒グラフとしても確認できる。

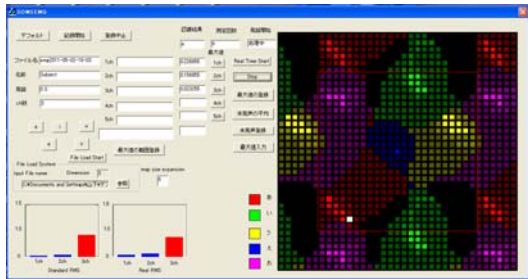


図8 SOMを用いた訓練システムの外観

無発声動作による発話認識を経験していない被験者10名に対して、5名ずつの二つのグループに分けて、一方には従来のコンピュータ合成の音声と表示される文字によって認識結果が確認できるシステムを使用し、もう一方には提案した訓練システムを使用し、試行回数と認識率の関係を求めた。発話は母音をそれぞれ20回連続して発話し認識率を求めた。実験では、「認識実験」、「2分間の休憩」、「5分間の自由な練習」、「認識実験」、「2分間の休憩」、「5分間の自由な練習」、「認識実験」の手順で、2回の練習時間と、3回の認識実験を行った。これを4セット、二日間の午前・午後に分けて実施し、合計12回の認識実験の認識率の推移を調べた。10名の被験者で、初回の5母音の平均認識率は60%前後と大きな差はみられなかった。5名分の母音の平均認識率をそれぞれ求めグラフにプロットしたものを図9に示す。

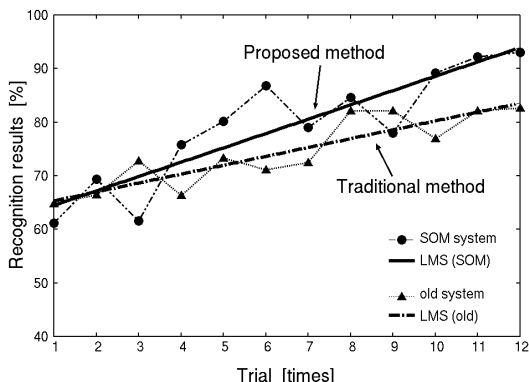


図9 試行回数と母音認識率（平均）の関係

認識率の推移を最小二乗法によって求めた直線を重ねて示すと、訓練システムを使用することによって、従来のシステムに比べて、短い試行回数で認識率が上昇していること

が分かり、提案した訓練システムの有効性を示すことができた。提案システムを使用した被験者では、最終的にほぼ90%の認識率を得ることができるようになっている。

(3) 一部の子音を含む発話認識

安定した無発声発話動作が可能になった被験者について、/ma/, /wa/, /ta/, /na/などの子音を含めた発話に関する検討を行った。子音を含む発話認識においては、母音の筋電パターンに比べて、わずかな振幅の差異や、時間遅れが見られることを予備実験で確認していた。しかし、これらの特徴は、母音だけでの特徴の差異に比べて小さいものであるため、安定した筋動作の獲得は不可欠である。十分に無発声動作の経験を積んだ被験者1名において、/ma/や/wa/の発話において、口輪筋(CH1)周辺筋肉が、/a/に比べてわずかに大きく活動することを利用して、CH1の閾値を2段階に設定し、/ma/あるいは/wa/を認識できるようにして、認識実験を行った。

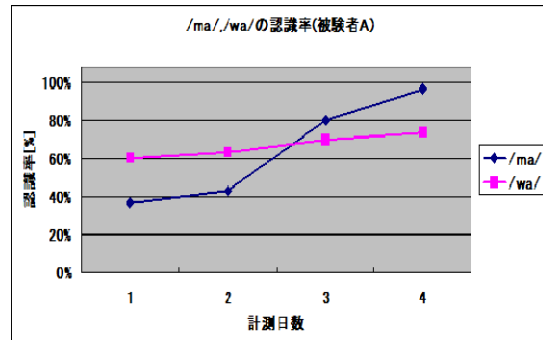


図10 /ma/・/wa/の認識率の推移

図10に示すように、試行回数を重ねる毎に、認識率も向上しており、最終的には/mā/については、95%以上の認識率を得ることができた。しかし、現状では、/ma/と/wa/を区別することはできていない。

同様に、/ta/, /na/, /ya/のように、発話時に舌を大きく動かすことから、胸骨舌骨筋(CH3)周辺に、/a/と比較して筋活動に時間遅れが見られることを確認し、時間差を利用した認識実験を行った。

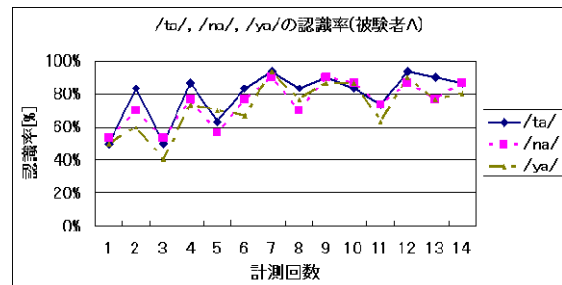


図11 /ta/, /na/, /ya/の認識率の推移

図11に試行回数と認識率の推移を示す。

認識率の変動はあるものの、試行回数を重ねる毎に、しだいに高い認識率で安定してきていることがわかる。また、ここでも、/ta/, /na/, /ya/をそれぞれ区別することはできていない。

このように、一部の子音を含む発話認識に関する検討を行い、1名の被験者での検討ではあるが、予備実験で得られた別の被験者との傾向が一致しており、子音を含む音節の認識に一部成功したと言える。今後は、被験者の数を増やして検討を進める予定である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計1件)

① 福本尚生, 野口祐介, 相知政司, 古川達也 : “表面筋電信号による無発声母音認識システムにおける自己組織化マップを用いたパターン解析”, 電気学会論文誌 C, 査読有, Vol.129, No.10, pp.1785-1791 (2009)

[学会発表] (計4件)

① 山下明大, 倉富勲, 福本尚生, 古川達也, 伊藤秀昭, 和久屋寛, 相知政司 : “子音の無発声時における筋電信号の遅れに関する検討”, 日本生体医工学会九州支部学術講演会論文集, 27, p.49, 2011年1月8日, 九州大学

② 倉富勲, 山下明大, 福本尚生, 古川達也, 伊藤秀昭, 和久屋寛, 相知政司 : “自己組織化マップを用いた無発声母音認識のための訓練システム”, 電気学会計測技術委員会主催第1回学生研究発表会予稿集, p.1, 2010年11月18日, 佐賀大学

③ 山下明大, 福本尚生, 古川達也, 和久屋寛, 伊藤秀昭, 相知政司 : “表面筋電信号による無発声認識システムを用いた発話認識の検討”, 第63回電気関係学会九州支部連合大会, 01-2P-14, CD-ROM, 2010年9月26日, 九州産業大学

④ 福本尚生, 相知政司, 古川達也 : “無発声発話時における口唇周辺表面筋電信号の自己組織化マップによる解析”, 第48回日本生体医工学会大会 生体医工学第47巻特別号 VOL.47, No.Suppl.1, 2009年4月24日, 東京都

6. 研究組織

(1) 研究代表者

福本 尚生 (FUKUMOTO HISAO)
佐賀大学・工学系研究科・助教
研究者番号 : 60346872