

筋電位信号で操作可能なテキスト入力システム

辻 敏夫*¹ 重藤 元暢*¹ 岡本 勝*² 杉山 利明*¹ 柴 建次*¹

A Text Input System manipulated by EMG Signals

Toshio Tsuji*¹, Motonobu Shigeto*¹, Masaru Okamoto*², Toshiaki Sugiyama*¹ and Kenji Shiba*¹

Abstract – Electromyographic (EMG) signals contain useful information on muscular activities and can be measured noninvasively using surface electrodes. This paper proposes a novel text input system manipulated by EMG signals, that reflect muscular activities associated with vocalizations, for the physically challenged. In this system, first, user's intention of motion can be estimated with a neural network from EMG signals. Then, the estimated motions are corresponding to operating commands of the system, so that users can input arbitrary characters using their body motions. The differences in the measured EMG signals among individuals can be compensated by using adaptive learning of the neural network, and the operation mode of the system can be changed in accordance with user's ability. In addition, words are predicted from some sequences of characters using the hidden Markov model, so the manipulation time required for input of sentences is considerably reduced. The effectiveness of the proposed system is demonstrated with text input experiments by a cervical spine injury patient who cannot operate a keyboard. It is shown from the experimental results that the proposed methods can be used by the physically challenged.

Keywords : Text Input, EMG signals, Probabilistic Neural Network, Hidden Markov Model

1. はじめに

近年、急激な情報化社会の発展により、広く一般の家庭までPCなどの情報機器が普及し、専門家だけでなく、誰もがメールやインターネットを利用できる環境が整いつつある。そのため、身体障害者もこれらを使用する機会が増えてくると考えられるが、現在のPCのテキスト入力装置として使用されているキーボードは、身体障害者にとっては操作が困難であり、症状によっては使用できない場合がある。もし、身体障害者が自由にテキスト入力を行いメールなどを扱えるようになれば、コミュニケーションの幅が広がり生活の向上につながると考えられるが、そのためには誰もが容易に使用可能な新しいテキスト入力装置の開発が必要である。

これまでに、キーボードに代わるさまざまなテキスト入力装置が開発されてきた。Isokoskiは操作者の視線の動きをカメラを用いてとらえることで文字を選択、入力する研究を行っている^[1]。これは、表示画面の外枠の上下左右に四つの点を表示し、操作者が視線で追った順序によって入力する文字を決定する。また、IBM社のViaVoiceは操作者が話した言葉をテキスト入力するものであり^[2]、容易に使用できるため障害者

に限らず多くの人々が利用している。しかしながら、これらの装置は障害の部位や度合いによっては扱えない場合がある。一方、Trejoらは筋収縮時に発生する筋電位信号(以下、EMG信号)を利用した仮想キーボードの研究を行っている^[3]。この方法では前腕から計測したEMG信号を利用し、学習識別を行うことで指の動作を推定している。EMG信号は筋が収縮する箇所さえあれば計測可能なため、残存機能をうまく活用できるといった利点がある。しかしながら、現段階ではテンキーのみの識別しか行えないため、任意の文章入力はできない。さらに、福田らはボコーダシステムと呼ばれる代用発声システムを開発した^[4]。このシステムは、顔から計測したEMG信号を用いて識別した母音から、隠れマルコフモデル(Hidden Markov Model以下、HMM)により単語推定を行い、音声を再生するシステムである。このシステムは、入力された6語音から指定された単語をHMMを用いて推定できる。しかしながら、母音から単語推定を行うため単語の数が増えるにつれて単語推定が難しくなるため、任意の文章入力は困難である。同様に、真鍋らは顔のEMG信号から母音の識別を試みている^[5]。

本論文では、EMG信号を利用したテキスト入力システムを提案する。このシステムは、EMG信号から操作者の動作を推定し、その動作を利用して操作画面上に配置された仮名文字を選択決定し、任意の文字列の生成を可能としている。また、全ての文字を入力す

*1: 広島大学大学院 工学研究科

*2: 広島市立大学 情報科学部

*1: Graduate school of Engineering, Hiroshima University

*2: Faculty of Information Sciences, Hiroshima City University

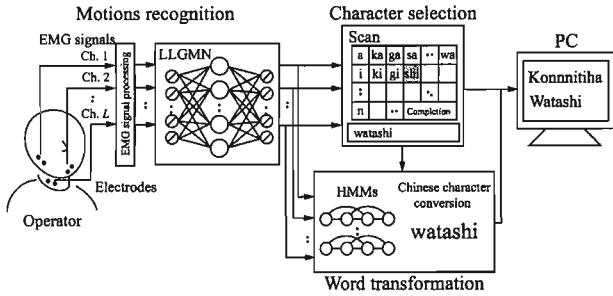


図1 テキスト入力システム

Fig.1 Overview of the text input system

る前に単語を推定する予測変換機能や、母音列から単語推定を行うボコーダシステムの一部の機能を導入することで、文章を効率よく生成できる。

以下、2. では構築したシステムの構成について述べ、3. と4. の実験でシステムの有効性を検証する。

2. システム構成

Fig. 1 に、開発したテキスト入力システムの構成を示す。本システムは、身体のさまざまな部位から計測した EMG 信号の振幅情報を利用し、操作者の意図する動作の識別を確率ニューラルネットにより行う。そして、識別した動作を用いて文字を選択し、文字列を作成する。さらに、作成した文字列を必要に応じて漢字変換、単語推定した結果の文章を PC のアプリケーションにテキストデータとして出力する。

以下、動作の識別を行う動作推定部、文字列を作成する文字選択部、漢字変換または単語推定を行う単語推定部についてそれぞれ説明する。

2.1 動作推定部

2.1.1 特徴抽出処理

双極表面電極から測定した L 対分の EMG 信号を A/D 変換 (サンプリング周波数: f_s [Hz]) し、チャンネルごとに全波整流を行う。そして、2 次のデジタルバターワースフィルタ (カットオフ周波数: f_c [Hz]) により平滑化処理を施し、得られる時系列信号を $E_l(t)$ ($l = 1, 2, \dots, L$) とする。安静時に発生する微小な EMG 信号の影響をなくすため、 $E_l(t)$ から安静時における $E_l(t)$ の時間平均である \bar{E}_l^{st} を引き、各チャンネルごとに規格化したものを $\alpha_l(t)$ と定義する。

$$\alpha_l(t) = \frac{E_l(t) - \bar{E}_l^{st}}{E_l^{max} - \bar{E}_l^{st}} \quad (1)$$

$(l = 1, 2, \dots, L)$

ここで、 E_l^{max} は事前に計測した各チャンネルの最大随意収縮時の値を表す。その後、全チャンネルの和が 1 となるように正規化したものを、特徴パターンベクトル $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_L(t)]^T$ としてパターン識別の入力ベクトルに用いる。

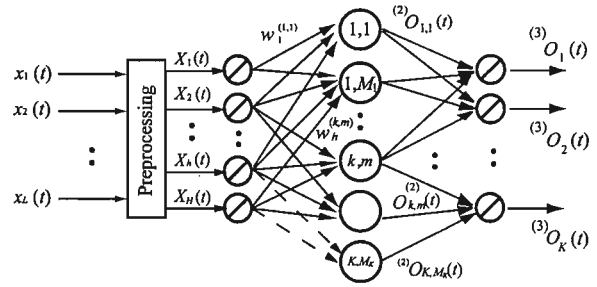


図2 LLGMN の構造

Fig.2 Structure of LLGMN

$$x_l(t) = \frac{\alpha_l(t)}{\sum_{l'=1}^L \alpha_{l'}(t)} \quad (2)$$

また、次式で示す $F_{EMG}(t)$ を本論文では筋力情報と定義し、筋力情報 $F_{EMG}(t)$ を用いて動作決定の開始と終了の判断を行う。

$$F_{EMG}(t) = \frac{1}{L} \sum_{l=1}^L \alpha_l(t) \quad (3)$$

2.1.2 パターン識別

パターン識別には辻らが提案した Log-Linearized Gaussian Mixture Network (以下、LLGMN と略記 [6]) を用いる。LLGMN は混合ガウス分布モデル (Gaussian mixture model: 以下、GMM と略記) を対数線形化してネットワークに展開したもので、サンプルデータが従う統計分布を学習的に推定し、入力パターンの各出力クラスに対する事後確率を出力することができる。LLGMN の学習能力により、個人差による EMG 信号の違い、電極位置のずれなどによる EMG 信号の変化などに対応できる。

LLGMN の構造を Fig. 2 に示す。まず GMM の各コンポーネントに対応する正規分布をニューラルネットワークの重み係数として表現するために、時刻 t における入力ベクトル $\mathbf{x}(t)$ を次式のように変数変換する。

$$\mathbf{X}(t) = [1, \mathbf{x}(t)^T, x_1(t)^2, x_1(t)x_2(t), \dots, x_1(t)x_L(t), x_2(t)^2, x_2(t)x_3(t), \dots, x_2(t)x_L(t), \dots, x_L(t)^2]^T \quad (4)$$

ネットワークの第 1 層はこの生成された入力ベクトル $\mathbf{X}(t)$ の次元数 $H = 1 + L(L+3)/2$ に合わせて H 個のユニットから構成される。各ユニットは恒等関数を入力関数として用いており、 $X_h(t)$ ($h = 1, 2, \dots, H$) を出力する。この第 1 層の出力に重み $w_h^{(k,m)}$ が掛け合わせられ第 2 層に伝えられる。第 1 層の出力を $^{(1)}O_h(t)$ 、第 2 層のユニット $\{k, m\}$ ($k = 1, 2, \dots, K, m_k = 1, 2, \dots, M_k$) への入力を $^{(2)}I_{k,m}(t)$ とし、出力を $^{(2)}O_{k,m}(t)$ とすると、

$${}^{(2)}I_{k,m}(t) = \sum_{h=1}^H {}^{(1)}O_h(t)w_h^{(k,m)} \quad (5)$$

$${}^{(2)}O_{k,m}(t) = \frac{\exp[{}^{(2)}I_{k,m}(t)]}{\sum_{k'=1}^K \sum_{m'=1}^{M_{k'}} \exp[{}^{(2)}I_{k',m'}(t)]} \quad (6)$$

となる。ただし、 $w_h^{(K,M_K)} = 0$ ($h = 1, 2, \dots, H$) で、 K は対象とする動作の数、 M_k は動作 k に属するガウス分布のコンポーネント数を表す。さらに第3層のユニット k は第2層の M_k 個のユニットと結合しており、その入出力関係は、

$${}^{(3)}I_k(t) = \sum_{m=1}^{M_k} {}^{(2)}O_{k,m}(t) \quad (7)$$

$${}^{(3)}O_k(t) = {}^{(3)}I_k(t) \quad (8)$$

と表される。この第3層のユニットの出力値は、各動作に対する事後確率を与える。

動作決定には、EMG 信号より計算される筋力情報 $F_{EMG}(t)$ 、および LLGMN の出力を使用する。まず、あらかじめ閾値 F_d を設定しておき、筋力情報 $F_{EMG}(t)$ が F_d を超えたときに識別を開始する。この際、誤識別を防止するために、LLGMN の出力のエントロピーを利用して識別判定処理を行う。エントロピーは、情報のあいまいさを表す評価値であり、この値を利用することで、あいまいな識別を防ぐことができる^[7]。LLGMN が出力する事後確率より、次式のエントロピーを計算する。

$$H(t) = - \sum_{k=1}^K {}^{(3)}O_k(t) \log {}^{(3)}O_k(t) \quad (9)$$

この $H(t)$ を設定した閾値 H_d と比較し、 $H(t) < H_d$ であれば、出力 ${}^{(3)}O_k(t)$ ($k = 1, \dots, K$) が最大となるユニットに対応する動作を識別結果とする。 $H(t) \geq H_d$ であれば、あいまいな識別として識別を保留する。

2.2 文字選択部

文字選択部では、動作推定部で識別した動作を利用し、文字の選択を繰り返すことで任意の文字列を入力することができる。

Fig. 3 に文字選択画面を示す。陰影部はカーソルを示しており、本システムではこのカーソルが自動的に遷移し、決定に対応する動作を行うことでカーソルの示す文字を選択できる。さらに、句読点や delete キーなども配置しており、選択可能である。

操作者が随意的に制御できる動作数は、障害の部位や度合いによって異なる。本システムでは、識別動作数に応じて、以下に示す3つの動作モードを用意した。

(1) 1動作モード

1動作モードでは、動作をスイッチとして扱い文字入力を行う。まず、カーソルが文字選択画面の第1行を

		Direction of cursor movement																
		→																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
a	,	ka	ga	sa	za	ta		da	na	ha	ba	pa	ma	ya	ya	ra	wa	
i	,	ki	gi	shi	zi	ti		di	ni	hi	bi	pi	mi			ri		
u	,	ku	gu	su	zu	tu	tu	du	nu	hu	bu	pu	mu	yu	yu	ru		
e	,	ke	ge	se	ze	te		de	ne	he	be	pe	me			re		
o	,	ko	go	so	zo	te		do	no	ho	bo	po	mo	yo	yo	ro	wo	
n	,	.	Back Space	All Delete	Change	Kanji	Enter	①	②	③	④	⑤	⑥	⑦	⑧	Completion		

図3 文字選択画面

Fig. 3 An example of the character selection screen

左から右へ順にスキャンしていく。カーソルが右端まで移動した場合には、左端に戻りスキャンを続行する。そして、動作を入力するとそのカーソル位置で上下方向の移動に切り替わり、その列を上から下へスキャンしていく。なお、下端まで移動した場合には、上端に戻りスキャンを続ける。上下方向のスキャン中、選択したい文字にカーソルが移動した際に再度動作を行うことで、カーソルで選択している文字を決定する。

(2) 2動作モード

カーソルは第1行を左から右へ順にスキャンしていく。1つの動作を入力することでスキャン行を1行ずつ下に変更することができる。また、もう1つの動作の入力によりカーソルが選択している文字を決定する。

(3) 多動作モード

3動作以上の動作識別が可能な場合は、2動作モードにおける操作に加えて、動作数に応じてカーソルの移動方向の変更や、Back Space キー、Enter キーなどのコマンドを追加できる。そのため、使用者の識別可能な動作数が増えるにつれ、より効率的な文字選択が可能となる。

また、母音“あ”、“い”、“う”、“え”、“お”と撥音“ん”の計6語音発声時のEMG信号パターンのように6動作の識別が可能な場合は、以下に示すアルゴリズムで文字入力可能である。

- Fig. 4 に示すように、動作時に(2)式に示す筋力情報 $F_{EMG}(t)$ が閾値 F'_d ($F'_d > F_d$) を越えた区間の時間長を t' と定義する。
- t' と T' を比較し、 $t' \geq T'$ であれば識別結果自体を選択文字とし、1に戻る。 $t' < T'$ であれば識別結果に対応する行をカーソルが移動する。
- 2と同じ動作を、 $F'_d > F_{EMG}(t) \geq F_d$ の範囲で再度行うことによってカーソルの移動が停止し、カーソル位置に対応する文字が選択される。異なる動作を行った場合には、その動作に対応する行にカーソルを移動する。またカーソルの移動中においても、 $t' \geq T'$ を満たせば識別結果を選択文字とする。

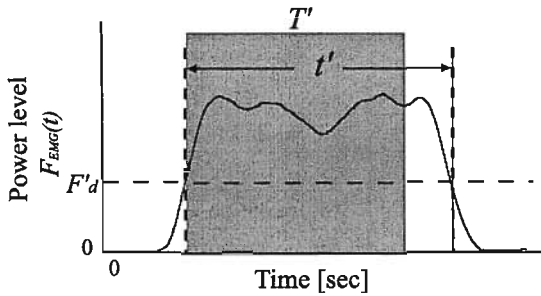


図4 t' の定義
Fig. 4 Definition of the motion time t'

このように、“あ”、“い”、“う”、“え”、“お”、“ん”の発声時の動作は、スイッチ操作として利用するとともに、直接的な文字入力としても利用可能である。

また、コマンド決定時における筋力情報 $F_{EMG}(t)$ を利用することにより、4段階のカーソル移動速度の調整が可能である。4段階のカーソル移動における1文字あたりの遷移時間を、それぞれ t_{s1} , t_{s2} , t_{s3} , t_{s4} [sec] ($t_{s1} > t_{s2} > t_{s3} > t_{s4}$) と定義する。そして、あらかじめ設定した筋力情報の閾値 F_{s1} , F_{s2} , F_{s3} ($F_{s1} < F_{s2} < F_{s3}$) に対して、 $F_{EMG}(t) < F_{s1}$ のときは t_{s1} , $F_{s1} \leq F_{EMG}(t) < F_{s2}$ のときは t_{s2} , $F_{s2} \leq F_{EMG}(t) < F_{s3}$ のときは t_{s3} , $F_{s3} \leq F_{EMG}(t)$ のときは t_{s4} とする。使用者は文字の決定や、スキャン行の変更などのコマンド決定時における $F_{EMG}(t)$ を大きくすることで、カーソル移動速度を速く、 $F_{EMG}(t)$ を小さくすることで、カーソル移動速度を遅くすることができる。これにより、使用者はシステム使用時の体調や慣れに合わせてカーソル移動速度を変更することができる。

以上のように、本システムでは操作可能な動作数に応じて動作モードを切り替え、任意の文字列を生成することができる。

2.3 単語推定部

単語推定部では文字選択部で生成した文字列の漢字変換や、次に入力される文字列の予測を行う。

漢字変換では漢字と読みに対応する仮名のデータベースに基づいて、入力と完全に一致する漢字への変換を行う。

予測変換では文字列から単語を推定するために隠れマルコフモデル^[8] (Hidden Markov Model 以下 HMM と略記) を用いる。本システムでは単語に対応する HMM を用意し、各モデルに文字列を対応付ける。HMM は文字列データをモデル化することが可能で、文字の入力毎にすべてのモデルにおいて事後確率を算出する。そのため、入力毎に意図する単語の確率が高くなり、文字列を全て入力しなくても予測変換が可能になる。文字列 $S_1 S_2 \dots S_N$ ($S_n \in \{\text{あ, い, } \dots, \text{ん}\}$) に対して、HMM によって推定された事後確率を

$P(n|S_1 S_2 \dots S_N)$ ($i = 1, 2, \dots, M$) とすると、本システムでは全モデルの確率の和が1となるように正規化を行った O_i を出力とする。

$$O_i = \frac{P(n|S_1 S_2 \dots S_N)}{\sum_{i=1}^M P(n|S_1 S_2 \dots S_N)} \quad (10)$$

ただし、 M は HMM のモデル数、 N は入力した文字数を表している。

HMM の学習には、単語に対応する仮名の文字列データを用いる。また、6 語音を用いた入力では母音を直接入力できるため、母音列から構成される文字列も学習に用いる。さらに本システムでは母音を入力する以外にも、選択時のタイミングのずれ等による誤入力を考慮し、誤入力の可能性があるデータも HMM の学習に加える。この誤入力に対応する学習データは、カーソル移動による選択を考慮すると、各文字に対して上下または左右に隣接する文字となり自動的に決定できる。

以上の漢字変換候補や予測変換候補をディスプレイ上に表示し、Fig. 3 の選択画面に候補選択を行うためのコマンドを追加することによって、同音異字単語なども容易に入力することができる。

3. 操作実験

3.1 実験条件

開発したシステムの有効性を検証するため実験を行った。被験者は健常な男性3名 (A: 21 歳, B: 24 歳, C: 23 歳) と頸髄損傷患者1名 (D: 51 歳) の計4名とした。なお、被験者 D は第5頸髄を損傷しており、機能レベルとしては C5 レベルの四肢麻痺による ADL 全介助レベルである。使用した電極は直径 10 [mm] の銀電極であり、電極数は5対 ($L = 5$) とした。電極装着位置は、EMG 信号の振幅が顕著に見られる部位である。Ch. 1: 口角下制筋, Ch. 2: 大頬骨筋, Ch. 3: 咬筋, Ch. 4: 顎二腹筋, Ch. 5: 下唇下制筋とした。EMG 信号の測定にはマルチテレメータ (NEC 三栄 MT-11) を使い、80 [dB] の利得で増幅した。また、 f_s は 1000 [Hz], f_c は 1 [Hz] とし、LLGMN の学習は実験直前に被験者ごとに行った。なお、被験者には事前に実験目的や内容を十分に説明し、同意を得た上で実験を行った。

3.2 6 語音の識別

まず、6 動作の識別実験を行い、識別可能な動作数を確認した。動作は、“あ”、“い”、“う”、“え”、“お”、“ん”の6語音発生時の口の形とし、被験者には、一定間隔で6動作を行うよう指示した。なお、被験者は識別結果がフィードバックされる画面を見ながら、識別練習を30分程度行った後に識別実験を行っ

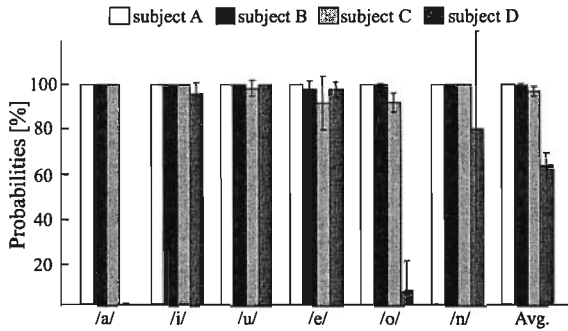


図5 6 語音の識別実験結果
Fig.5 Results of discrimination experiment

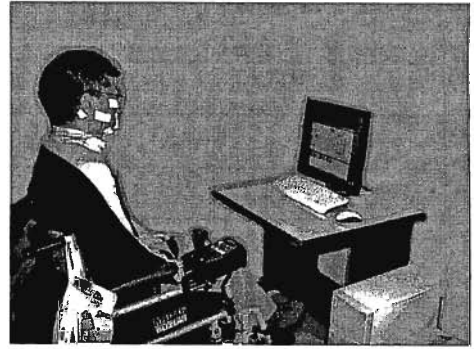


図6 テキスト入力システムの実験風景
Fig.6 Experimental view of the text input system

た。また、 H_d は0.3、 F_d は被験者A、B、Cにおいては0.15、被験者Dは0.2とし、試行数は5回とした。なお、提案法に含まれる各種のパラメータは、識別練習において各被験者に適した値を試行錯誤的に決定した。パラメータの決定法については、今後、検討していきたいと考えている。

各被験者の識別結果を Fig. 5 に示す。各値は5試行の識別率の平均値と標準偏差を示している。図より、被験者A、B、Cは6動作を精度よく識別できているが、被験者Dの結果では“あ”、“お”の識別率が低下し、“ん”の識別はばらつきが大きいことが確認できる。識別結果より、被験者A、B、Cは6動作、被験者Dは“い”、“う”、“え”の3動作が識別可能であることが明らかになった。

3.3 システム操作

3.2の識別結果を利用し、システムの操作を行った。22個の単語のHMMをBaum-Welch法を用いて作成し^[8]、その予測変換を可能にした。なお、 H_d と F_d は3.2と同様の値とした。

開発したシステムを用いて被験者Dがテキスト入力を行っている様子を Fig. 6 に示す。また、その時のディスプレイの一例を Fig. 7 に示す。Fig. 7より、“ん”行に配置されている番号(1~8)を選択することで、予測変換候補(Fig. 7(A))、漢字変換候補(Fig. 7(B))の中から望みの単語を入力可能である。なお、漢字変換、予測変換候補は、“ん”行に配置されている切替コマンドを利用することで切替が可能である。

被験者Aの6動作による結果を Fig. 8 に示す。図は上から順に、5チャンネル分のEMG信号、筋力情報 $F_{EMG}(t)$ 、カーソルの横方向の位置、カーソルの縦方向の位置、動作の識別結果を表している。なお、カーソルの横方向の位置は Fig. 3の番号(1~17)と対応している。縦方向の破線は文字の選択及びコマンドの決定時刻、横方向の破線は F'_d 、陰影をつけた区間は T' を表している。また F'_d は0.30、 T' は0.50 [sec]とした。

実験では、“簡単で楽しい”という文章を生成する

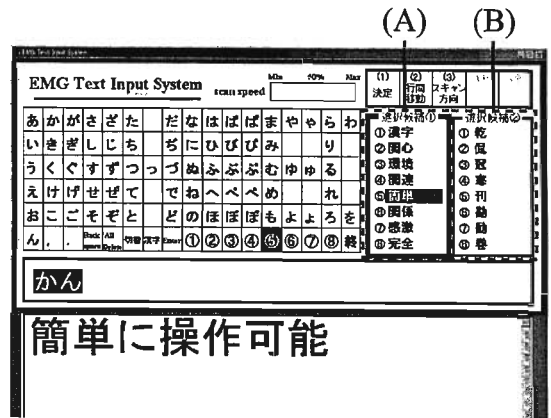


図7 テキスト入力システムの画面
Fig.7 Display of the text input system

よう被験者に指示した。Fig. 8では、“簡単”と入力する際、まず“あ”の発声動作を行うことで、“あ”行のスキャンが開始し、スキャン中に再度“あ”の動作を行うことで、“か”の文字を選択している。次に“ん”の入力では、 $F_{EMG}(t)$ が F'_d を越えた区間が T' 以上であるため、識別結果から直接文字を決定していることが分かる。そして、次に“た”の入力では、タイミングのずれにより“ざ”と誤入力をしてる。しかしながら、HMMの学習データとして誤入力を想定したデータを用いているため、(I)においてそのまま予測変換と対応するコマンドを選択することで、“簡単”と予測変換を行っている。また、“楽しい”と入力する際、“あおい”と母音を入力し、(II)において予測変換を行っているため、カーソル移動によって文字を選択する場合と比較し、少ない入力時間で単語が入力できた。そして、文章生成が終了した後、(III)においてPCのアプリケーションにテキストとして出力した。以上より6動作を利用することで、語音の直接入力と、カーソルで選択する文字入力を使い分け、さらに予測変換を用いて任意の文字入力が行えることが確認できる。

また、Fig. 8の0~21 [sec]の部分における筋力情報 $F_{EMG}(t)$ とカーソルの横方向の位置のグラフを Fig.

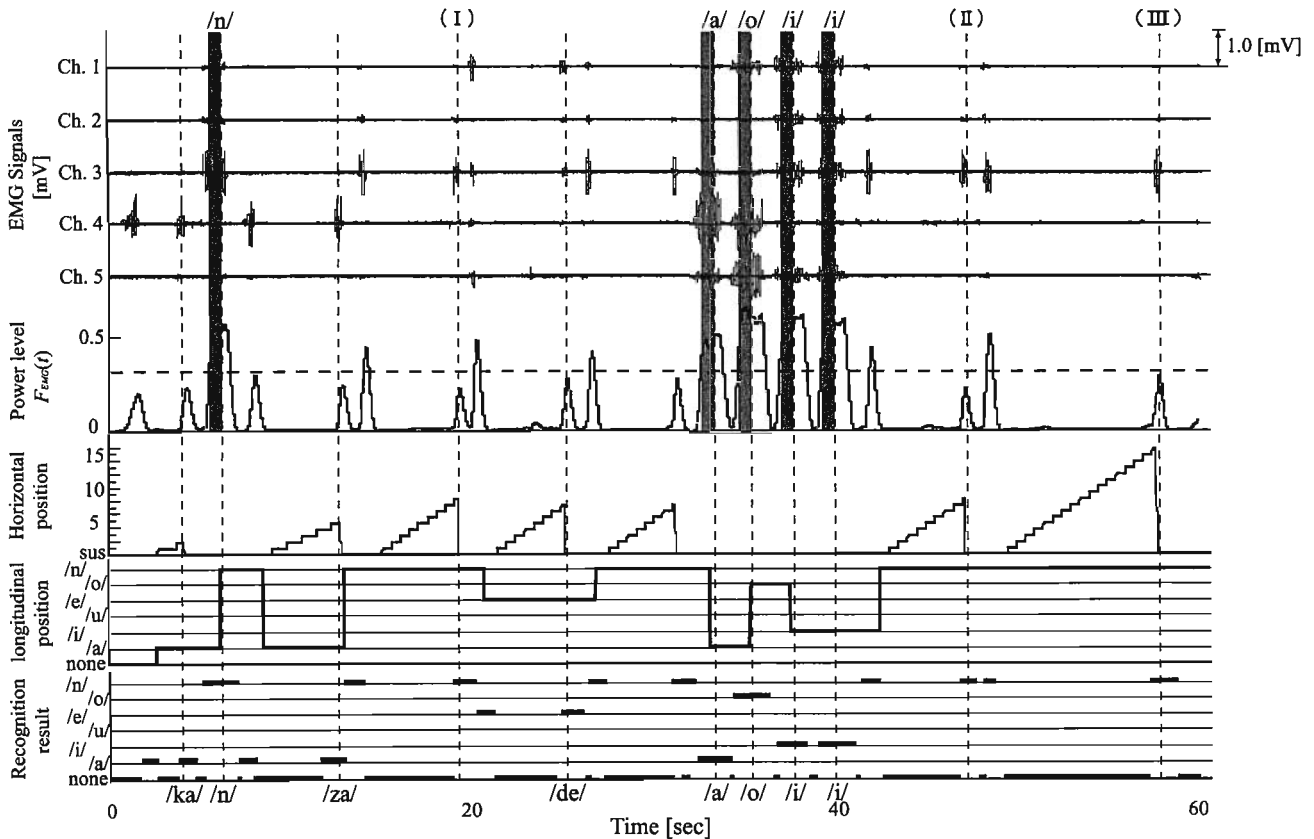


図8 被験者 A によるシステム操作実験の結果
Fig. 8 An example of the operation with subject A

9に示す。図中の縦方向の破線は文字の選択及びコマンドの決定時刻を表している。横方向の破線は閾値 F_{s1} , F_{s2} , F_{s3} を表しており、それぞれ 0.2, 0.3, 0.4 とした。また、カーソルの1文字あたりの遷移時間 t_{s1} , t_{s2} , t_{s3} , t_{s4} はそれぞれ 1.00 [sec], 0.83 [sec], 0.67 [sec], 0.50 [sec] と設定した。なお、これらの値は試行錯誤的に決定している。図の●はスキャン開始時における $F_{EMG}(t)$ の最大値を表しており、陰影をつけた区間は●の値によって決定した t_{sd} ($d = 1, 2, 3, 4$) を表している。まず、“か”を入力する際には、カーソルの初期位置からの距離が近いため、弱く発声動作を行い、 $F_{EMG}(t)$ を小さくすることで、遷移時間を t_{s1} に調節している。また、(I)の時は距離が遠いため、強く発声動作を行うことで $F_{EMG}(t)$ を大きくし、 t_{s4} に調節していることが分かる。以上より、選択する文字に応じて $F_{EMG}(t)$ の値を変え、カーソルの移動速度を調節できていることが確認できる。

次に Fig. 10 に被験者 D が“簡単で楽しい”という文章を入力した際の“楽しい”の部分の実験結果を示す。被験者 D は“い”，“う”，“え”の3動作を利用してシステムの操作を行った。なお、3動作モードにおけるコマンドは“文字の決定”，“スキャン行の移動”，“スキャン方向の反転”とし，“い”，“う”，“え”と順にコマンドを対応させた。Fig. 10では、ま

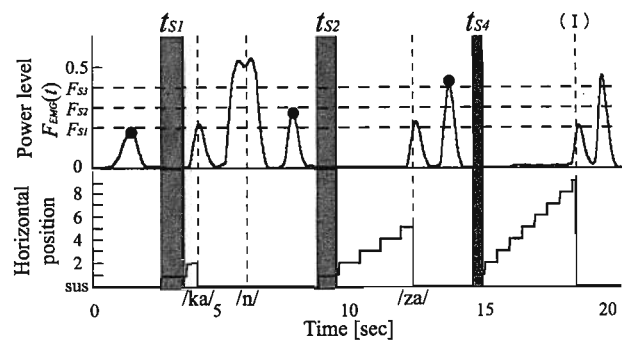


図9 被験者 A によるシステム操作実験の結果
Fig. 9 An example of the operation with subject A

ずカーソルを移動させ“たの”と入力し、その後カーソルの移動方向を変えて、(I)においてコマンドを選択することで“楽しい”と予測変換を行っている。また、文章の入力時間は 136.43 [sec] であった。これに対し被験者 A, B, C において、“あ”，“い”，“う”の3動作を用い、被験者 D と同条件で実験を行ったところ、5 試行分の平均がそれぞれ 70.37 ± 3.09 [sec], 79.70 ± 8.36 [sec], 96.13 ± 10.13 [sec] であった。ここで、被験者 D は事前にはほとんど練習を行わなかったため、操作時間が長くなったと考えられる。被験者 A, B, C に関しては事前にある程度練習を行った。そのため、ある程度練習を行うことで操作時間を短縮できる可能性がある。以上より、重度の身体障害者も随意

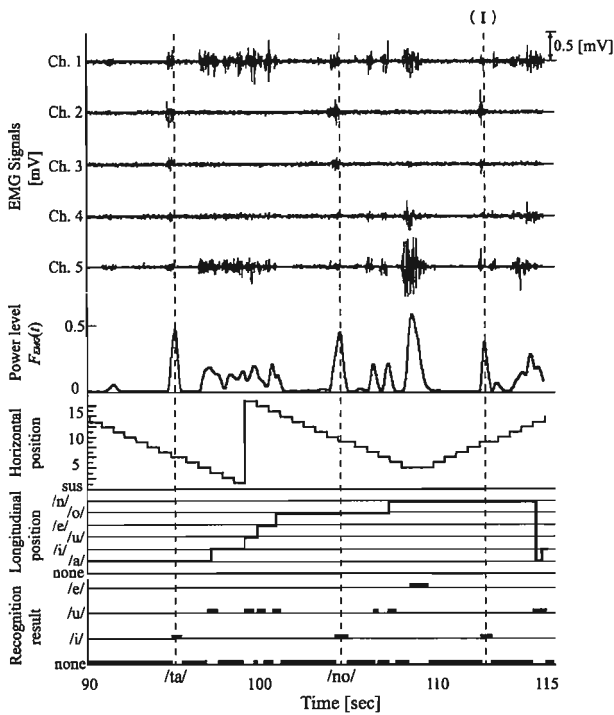


図 10 被験者 D によるシステム操作実験の結果
Fig. 10 An example of the operation with subject D

的に実行可能な動作数で文章生成が行えることが確認できた。

4. 予測変換の効果

本システムの予測変換の効果を検証するため、6 動作を用いて単語の入力時間の比較を行った。実験では、(1) 予測変換を用いない場合、(2) 文字列から予測変換を行う場合、(3) 母音列から予測変換を行う場合の 3 つの条件を対象とした。なお、実験条件は 3.1 と同様とした。

比較する単語は“私”，“楽しい”，“おはよう”，“非常に”の 4 つとし、(1) の場合は予測変換をせずにカーソルを動かして文字を入力し、完全一致の漢字変換を行うことで単語入力を行うよう指示した。また、(2) の場合はカーソル移動によって、それぞれ“わた”，“たの”，“おは”，“ひじょ”，と入力した後、予測変換するよう指示した。(3) の場合はそれぞれの単語の母音列を入力し、その後予測変換するよう指示した。被験者は A, B, C の 3 名で、試行回数は 5 回とした。ただし、誤入力は考慮せず、予測変換のみの効果を評価対象とした。またカーソルの 1 文字あたりの移動時間は、事前に予備実験を行い、0.67 [sec] とした。

実験結果を Fig. 11 に示す。各グラフは 5 試行の平均と標準偏差を示している。図より、どの単語においても (3) の場合が最も早く入力できており、入力時間もほぼ一定であることが確認できる。また (2) の場合も (1) と比べて入力時間が短縮されていることが分か

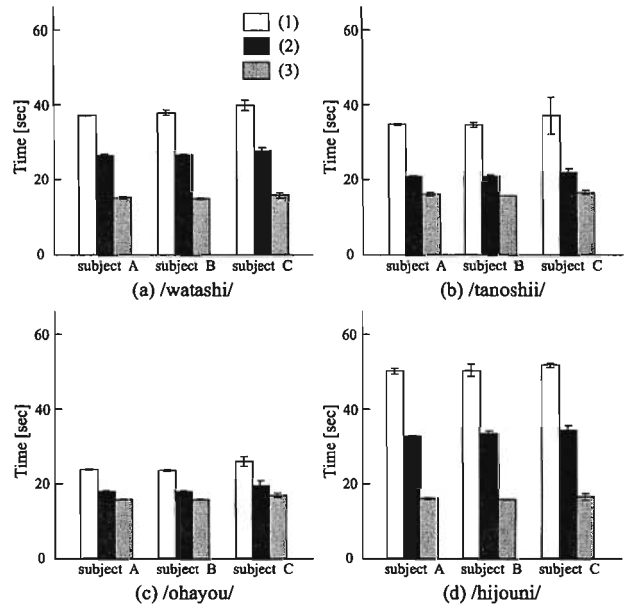


図 11 単語の入力時間
Fig. 11 Input time of words

る。“私”においては Fig. 3 より、“わ”の文字がカーソルの初期位置から遠いことから、(3) の場合においてより早い入力が行えたと考えられる。“非常に”の場合も同様のことが言える。“楽しい”の場合では“た”と“の”の文字は比較的カーソルの初期位置から近いいため、(2) と (3) の場合であまり差が見られなかった。しかしながら、(1) と比較すると入力時間の短縮が確認できた。“おはよう”という仮名で構成されている単語においては、どの場合においても大きな差は見られなかったが、(1) と比較すると、(2), (3) の場合の方が入力時間が短い。以上より、入力時間は単語の文字数や文字位置によって異なるが、予測変換を行わない場合と比べて、文字列から予測変換を行ったほうがより短いことが確認できた。さらに、母音列から予測変換を行うと、単語にかかわらず安定して速く入力ができ、特にカーソルの初期位置から遠い文字列で構成される単語において効果的であることが確認できた。しかしながら、HMM は単語数が増加するにつれて推定が難しくなるため、全ての単語を予測変換可能とすることは困難である。そのため、使用者がよく扱う単語を予測変換として用い、それ以外の単語は漢字変換などを用いることで効率のよい文章生成が可能になると考えられる。

5. まとめ

本論文では、キーボード操作が困難である身体障害者のための新しいテキスト入力装置として、EMG 信号を利用したテキスト入力システムを提案した。提案システムは操作者の状態に合わせた動作数で操作方法を変更することができ、操作画面上に配置されている文字を選択することで任意の文字列が作成可能である。

(2006年10月31日受付, 2007年2月13日再受付)

著者紹介

さらにHMMを用いることで、作成した文字列から単語の予測変換が可能で、素早い文章生成が行える。そのため、身体障害者だけでなく、健常者もウェアラブルコンピュータなどに使用できる可能性がある。

実験では、キーボード操作が行えない頸髄損傷患者が識別可能な動作数でテキスト入力が行えることを確認した。また、単語の入力時間を比較し予測変換の有効性を示した。

今後は予測変換可能な単語の増加を検討する必要がある。入力された単語の履歴を取り、頻繁に使用される単語に対して自動的にHMMのモデルを追加するアルゴリズムを構築することで、より使用者に合わせた使用が可能になると考えられる。また、操作画面の文字列やコマンドの配置を操作者に合わせて変更可能とすることで、より効率的な操作を実現していく予定である。

最後に、本研究を通じて実験にご協力して頂いた小林 勝さんに深く感謝いたします。なお、本研究の一部は21世紀COEプログラム「超速ハイパーヒューマン技術が開く新世界」の一環として行われたもので、ここに改めて謝意を表します。

6. 参考文献

- [1] Isokoski, P.: Text input methods for eye trackers using off-screen targets; *Proceedings of Tracking Research & Applications Symposium*, pp.15-22 (2000).
- [2] IBM, 連続音声認識プログラム, <http://www.scansoft.co.jp/viavoice/>
- [3] Trejo, L.J., Wheeler, K.R., Jorgensen, C.C., Rosipal, R., Clanton, S.T., Matthews, B., Hibbs, A.D., Matthews, R., Krupka, M.: Multimodal neuroelectric interface development; *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol.11, No.2, pp.199-204 (2003).
- [4] 福田 修, 藤田 真治, 辻 敏夫: EMG信号を利用した代用発声システム; 電子情報通信学会論文誌, Vol.J88-D-I, No.1, pp.105-112 (2005).
- [5] 真鍋 宏幸, 平岩 明, 杉村 利明: 無発声音声認識: 筋電信号を用いた声を伴わない日本語5母音の認識; 電子情報通信学会論文誌, Vol.J88-D-II, No.9, pp.1909-1917 (2005).
- [6] Tsuji T., Fukuda, O., Ichinobe, H., Kaneko, M.: A log-linearized gaussian mixture network and its application to EEG pattern classification; *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Application and Reviews*, Vol.29, No.1, pp.60-72 (1999).
- [7] 辻 敏夫, 市延 弘行, 伊藤 宏司, 長町 三生: エントロピーを用いた誤差逆伝播型ニューラルネットによるEMGからの前腕動作の識別; 計測自動制御学会論文集, Vol.29, No.10, pp.1213-1220 (1993).
- [8] Rabiner, L.R.: A tutorial on hidden markov model and selected application in speech recognition; *Proceedings of the IEEE*, Vol.77, No.2, pp.257-286 (1989).

辻 敏夫



1985 広島大学大学院工学研究科博士課程前期修了。同年同大学工学部助手。1994 同助教授を経て、2002より同大学大学院工学研究科教授、現在に至る。工学博士。人間とロボットの運動制御、生体信号解析、ニューラルネット、ヒューマン・マシンシステムなどの研究に従事。IEEE, 日本人間工学会, バイオメカニズム学会, 日本機械学会, 電気学会などの会員。

重藤 元暢 (学生会員)



2006 広島大学工学部第2類卒業。現在、広島大学大学院工学研究科博士課程前期過程在学中。EMG信号を用いたヒューマンインタフェースの研究に従事。ヒューマンインタフェース学会学生会員。

岡本 勝



2004 広島大学大学院工学研究科複雑システム工学専攻博士課程前期修了。2006 (株)アイピーシー勤務。2007より広島市立大学情報科学部知能情報システム工学科助教。ニューラルネット, 知識工学などの研究に従事。計測自動制御学会の会員。

杉山 利明



2005 広島大学工学部第2類卒業。2007 広島大学大学院工学研究科博士課程前期修了。現在, (株)キヤノン勤務。在学中はEMG信号を用いたヒューマンインタフェースの研究に従事。

柴 建次



2000 東京理科大学大学院理工学研究科博士課程修了。博士(工学)。日本学術振興会特別研究員(PD), 東京大学大学院新領域創成科学研究科助手, 東京理科大学理工学部助手を経て, 広島大学大学院工学研究科複雑システム工学専攻助教授。医用生体電子工学の研究に従事。日本生体医工学会, 日本人工臓器学会, 電気学会, 日本機械学会, ライフサポート学会, 日本人間工学会, IEEEなどの会員。