

EMG 信号を利用した代用発声システム

福田 修[†] 藤田 真治^{††} 辻 敏夫^{††}

A Substitute Vocalization System Based on EMG Signals

Osamu FUKUDA[†], Shinji FUJITA^{††}, and Toshio TSUJI^{††}

あらまし 本論文では、発声機能障害者のための新しい意思伝達装置の開発を目的とし、頸部及び表情筋から計測した EMG 信号に基づく代用発声システムを提案する。このシステムは、まず EMG 信号から使用者の意図する語音を推定し、次にそれを連ねた語音列から意図する単語を推定するという 2 段階の処理を行う。語音、及び単語の推定には、統計構造を有するニューラルネットと隠れマルコフモデルを用いた。従来の電気式人工喉頭は、頸部に人工呼吸器を取り付けた際などに使用が困難となるが、本手法はそのような場合でも使用することが可能である。健常な男子大学生及び喉頭切除者の計 5 名による実験を行った結果、精度良く語音・単語を識別できることが示された。

キーワード 代用発声, EMG 信号, ニューラルネット, 隠れマルコフモデル

1. ま え が き

音声言語は我々の日常生活において、意思伝達を瞬時に行うことのできる優れたコミュニケーション手段である。しかしながら、厚生労働省の身体障害児・者実態調査（平成 13 年度 6 月実施）によれば、喉頭癌などが原因で喉頭を切除し、発声機能に障害をもつ人は、日本国内だけでも 2 万人以上が存在する [1]。また、世界全体では 60 万人以上が存在するといわれている [2]。また、このような障害者は高齢者に多いことから、その数は今後更に増えることも予想される。

喉頭発声に代わる新しい発声法を容易に習得し、明りょうな音声言語を再獲得することができれば、発声機能障害者にとって最大の希望となるであろう。健常者の喉頭発声は、呼吸を喉頭に送り出す肺、呼吸を音源に変換する声帯、及び喉頭から口唇までの構音器官の三つの協調動作によって生成されるが、喉頭切除者や発声機能障害者は、音源である声帯、舌や下顎等の

構音器官に十分な機能が得られず発声が困難となる。

これまでに、代用発声法に関する研究が数多く試みられてきた。例えば関らは、優れた代用発声法とは、発声のタイミング、強さ、ピッチ、持続時間などを自由に制御できるもので、両手が使えるなど日常生活に支障がないものと定義している。そして、従来の代用発声法の問題点を挙げるとともに、電気式人工喉頭制御を提案している [3]。この電気式人工喉頭は、習得が容易であること、呼吸を利用して抑揚のある自然な発声が行えることなどから、既に 1000 台以上が実際に利用されており、約 7 割の使用者から良好な反応が得られている [4]。しかしながら、喉頭切除者などが頸部に人工呼吸器を取り付けた場合には、装置を装着できないなどの問題が残された。

電気式人工喉頭以外の方法としては、舌や口唇の運動から語音識別を行い、音声を合成・再生する方法が研究されている [5], [6]。しかしながら、この運動をビジョンセンサなどにより計測する場合、語音によっては口の形状が無発声と同様になるため、高精度な識別は困難であった。また、操作者が意図する流ちょうな音声（タイミング、強さ、ピッチなど）を合成することも難しかった。これに対して、筋収縮に伴って生じる筋電位信号（以下 EMG 信号）を利用して流ちょうな発声を実現することが試みられている。例えば村木らは、表情筋、及び頸部の筋の EMG 信号から母音の

[†] 独立行政法人産業技術総合研究所人間福祉医工学研究部門，つくば市

Institute for Human Science and Biomedical Engineering, National Institute of Advanced Industrial Science and Technology, Tsukuba-shi, 305-8564 Japan

^{††} 広島大学大学院工学研究科，東広島市

Graduate School of Engineering, Hiroshima University, Higashi-hiroshima-shi, 739-8527 Japan

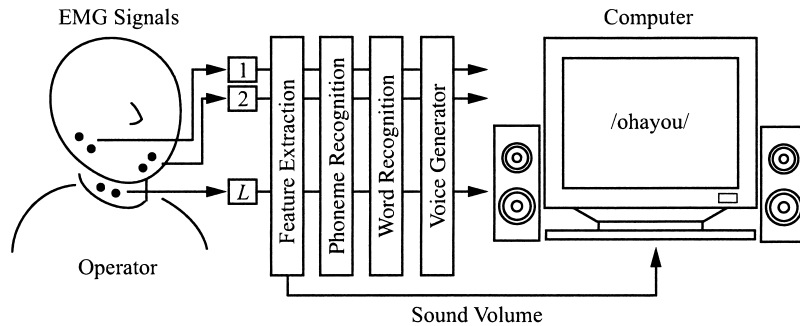


図 1 EMG 信号を用いた代用発声システム

Fig. 1 Substitute vocalization system based on the EMG signals.

識別を行っている [7], [8] . しかしながら, その識別精度は 60%程度であり, 安定した識別は実現されていない. 真鍋らは, バックプロパゲーション型ニューラルネットワークを用いて 5 母音の識別を行っている [9] . この研究では 3 チャネルの EMG 信号から, 90%以上の識別精度が実現されているが, 単語の識別などについては言及されていない.

本研究では, 発声機能障害者のための新しい意思伝達装置の開発を目的とし, 頸部及び表情筋の EMG 信号を利用した代用発声システムを提案する. 以下 2. では, 提案するシステムの概要について述べ, 3. ではその構成を説明する. 4. では健常者及び喉頭切除者による実験を行い, 提案手法の有効性を検証する. 5. にまとめを示す.

2. EMG 信号を利用した代用発声システム

言語による意思伝達は, 声の大きさや抑揚などが極めて重要な役割を果たしていることが明らかである. この際, 喉頭や口唇の周辺では, 発声に関与する筋が協調して機能している. 例えば, 大きな声で発声を行う場合には, 自然に口の形状を大きくし, 頸部を緊張させるように表情筋や頸部の筋が働く. 本研究では代用発声のための手段として, これらの筋が発生する EMG 信号に着目する. もし, この信号から発声しようとする単語, 声の大きさ, 抑揚などを推定することができれば, 自然な代用発声が可能である可能性がある. 我々は, これまでに EMG 信号からの筋力, 関節インピーダンス推定, ニューラルネットワークによる EMG パターンの識別, 及びこれらに基づく電動動力義手の制御に関する研究を行っており [10] ~ [12], これらの研究で得た知見を代用発声システムに導入することを試みる.

EMG 信号から発声情報を抽出するためのアプローチとして, 本論文では図 1 のような 4 段階の信号処理を考える. これらは, 1. EMG 信号からの特徴抽出処理, 2. EMG 特徴量からの語音識別処理, 3. 語音列からの単語識別処理, そして 4. 発声のための音声生成処理で構成される. EMG 特徴量からの語音識別処理には, 統計構造を内包したニューラルネットワークを用いる [13]. ここでは, ニューラルネットの適応学習機能により, EMG 特徴量に含まれる個人差を吸収することができる. また, ニューラルネットの識別結果を連ねた語音列から単語を識別する処理には, 隠れマルコフモデル [14] を利用する. この際, 識別対象の単語は, 日常生活において必要なものをあらかじめ候補として登録しておく. 隠れマルコフモデルは, 音声識別などによく用いられる手法であり, モデル化する対象の時間収縮にロバストであるなどの優れた特徴を有する. 音声の再生には, パソコン上で動作する音声合成ソフトウェアを用いる.

図 2 の写真は, (a) 頸部に呼吸用のチューブを付けた喉頭切除者の様子, 及び, (b) 開発したシステムを使用する様子を示すものである. 本システムは, 使用者の EMG 信号を利用して代用発声を実現するため, 頸部に人工呼吸器を取り付けた場合でも使用することが可能である.

3. システム構成

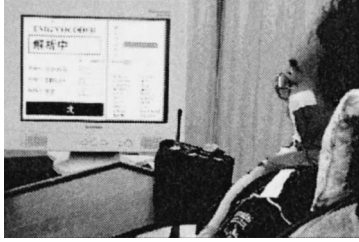
図 1 における特徴抽出処理, 語音識別処理, 単語識別処理, 及び音声生成処理について, 以下に詳細を説明する.

3.1 特徴抽出処理

EMG 信号の計測には生体アンプ (NEC 三栄 MT-11), 及び L チャネル分の表面皮膚電極 (GE mar-



(a) A laryngectomee



(b) Developed system

図 2 喉頭切除者と開発した代用発声システム

Fig. 2 Photos of a laryngectomee and the developed system.

quette 製 SEB120) を用いる。まず、電極から測定した EMG 信号を A-D 変換 (サンプリング周波数 1.0[KHz]) し、チャンネルごとに全波整流した後、2 次のデジタルパワースフィルタ (カットオフ周波数 1.0[Hz]) により平滑化する。そして、この時系列信号を $E_l(t)$ ($l = 1, \dots, L$) とする。

まず、発声の有無を判断するために、時刻 t における推定筋力 $F_{EMG}(t)$ を次式で定義する。

$$F_{EMG}(t) = \frac{1}{L} \sum_{l=1}^L \frac{E_l(t) - E_l^{st}}{E_l^{max} - E_l^{st}} \quad (1)$$

ここで、 E_l^{st} は安静時における $E_l(t)$ の時間平均、 E_l^{max} は実験前に計測した最大随意筋収縮時における時間平均を表す。この $F_{EMG}(t)$ と設定したしきい値 F_d を比較することで、語音識別を実施するタイミングを検出する。

ニューラルネットによる語音識別は、全チャンネルの和が 1 になるように正規化した特徴パターンベクトル $x(t) = [x_1(t), x_2(t), \dots, x_L(t)]^T \in \mathbb{R}^L$ を用いる。

$$x_l(t) = \frac{1}{L} \frac{E_l(t) - E_l^{st}}{F_{EMG}(t)} \quad (2)$$

このパターンベクトル $x(t)$ は、力情報を使って正規化されており、発声の大小に依存しない筋協調の割合を示している。

3.2 語音識別処理

語音識別には、辻らが提案した Log-Linearized Gaussian Mixture Network (以下、LLGMN と略記) [13] 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することができる。この学習能力は、個人差や電極位置、疲労や発汗などに伴う EMG 信号の変化を吸収し、精度の高い識別を実現する。

本研究では、識別対象の語音を母音“あ”、“い”、“う”、“え”、“お”及び、撥音“ん”の計 6 語音に限定し、子音については、それが属する母音 (例えば、“か” → “あ”、“き” → “い”) として識別するものとする。これは、EMG 信号だけから母音・子音のすべてを識別することは、現状では極めて困難なためである。つまり、操作者が“おはよう”と発声すると、システムは“おおおう”と識別する。

LLGMN の構造を表 1 に示す。まず、時刻 t における入力ベクトル $x(t)$ を次式のように変数変換する。

$$X(t) = [1, x(t)^T, x_1(t)^2, x_1(t)x_2(t), \dots, x_1(t)x_L(t), x_2(t)^2, x_2(t)x_3(t), \dots, x_2(t)x_L(t), \dots, x_L(t)^2]^T \quad (3)$$

この変換は、GMM の各コンポーネントに対応する正規分布を、この新たな入力ベクトル $X(t)$ とニューラルネットの重み係数の線形計算で表現するために行う。ネットワークの第 1 層は、 $X(t)$ 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出力関数として用いており、入力 $X_h(t)$ をそのまま出力する。この第 1 層の出力に重み $w_h^{(k,m)}$ が掛け合わされ第 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 $\{k, m\}$ への入力を $^{(2)}I_{k,m}(t)$ 、出力を $^{(2)}O_{k,m}(t)$ とすると、

$$^{(2)}I_{k,m}(t) = \sum_{h=1}^H ^{(1)}O_h(t) w_h^{(k,m)} \quad (4)$$

$$^{(2)}O_{k,m}(t) = \frac{\exp[^{(2)}I_{k,m}(t)]}{\sum_{k'=1}^K \sum_{m'=1}^{M_{k'}} \exp[^{(2)}I_{k',m'}(t)]} \quad (5)$$

となる。ただし、 $w_h^{(K, M_K)} = 0$ ($h = 1, 2, \dots, H$) で、

表 1 LLGMN の構造
Table 1 Structure of LLGMN.

The 1st layer	Number of units	H
	Input	$X_h(t)$
	Output	$^{(1)}O_h(t)$
	I/O function	Identity function
The 2nd layer	Number of units	$\sum_{k=1}^K M_k$
	Input	$^{(2)}I_{k,m}(t) = \sum_{h=1}^H {}^{(1)}O_h(t) w_h^{(k,m)}$
	Output	$^{(2)}O_{k,m}(t)$
	I/O function	Generalized sigmoid function
The 3rd layer	Number of units	K
	Input	$^{(3)}I_k(t) = \sum_{m=1}^{M_k} {}^{(2)}O_{k,m}(t)$
	Output	$^{(3)}Q_k(t)$
	I/O function	Identity function
Weight coefficients from 1st layer to 2nd layer		$w_h^{(k,m)}$
Weight coefficients from 2nd layer to 3rd layer		1

$k = 1, 2, \dots, K$, $m = 1, 2, \dots, M_k$ である． K は対象とする語音の数， M_k は語音 k に属するガウス分布のコンポーネント数を表す．更に第 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，

$$^{(3)}I_k(t) = \sum_{m=1}^{M_k} {}^{(2)}O_{k,m}(t) \quad (6)$$

$$^{(3)}O_k(t) = {}^{(3)}I_k(t) \quad (7)$$

と表される．この第 3 層のユニットの出力値は，各語音に対する事後確率を与える．

識別語音の判定は，EMG 信号から計算した推定筋力 $F_{EMG}(t)$ ，及び LLGMN の出力をもとに実施する．まず，先に定義した $F_{EMG}(t)$ を設定したしきい値 F_d と比較し，発声実施のタイミングを調べる．そして，発声実施と判定された場合に語音識別を実行する．この際，誤識別を抑制するために，LLGMN の出力のエントロピーを利用して誤識別回避処理を行う [15]．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．

3.3 単語識別処理

単語識別処理では，LLGMN が識別した語音列から使用者の意図する単語を識別する処理を行う．この処理には，隠れマルコフモデル (Hidden Markov Model : 以下，HMM と略記) [14] を利用する．

HMM は，使用する単語それぞれに応じて，“おはよう” → “おあおう” のモデル 1，“食べる” → “あえう” のモデル 2 などと単語数分を用意する．そして，例えば被験者が“おはよう”と発声すると，それが属する語音列“おあおう”を表現したモデル 1 が選択され，語音から単語を識別することが可能となる．

単語識別において，大きな問題となるのは，使用者が発声する際のばらつきである．例えば，我々が“おはよう”と発声する場合を考えると，“おはよ”と短く発声したり，“おはよおー”と伸ばして発声するなど，発声者や状況によって様々な特徴が考えられる．HMM の利点は，このように被験者の違いによる特徴を，あらかじめ多様な学習信号を使って動的確率モデルとして表現することにある．これにより，音声に時間方向の伸縮がある場合にも，高いロバスト性でクラスタリングすることができる．

なお，本論文では離散型 HMM を使用し，LLGMN が推定した語音シンボル列（例：こんにちはーおおんいいいいあ）をあらかじめモデル化する．この際，サンプルデータは各単語 50 サンプル，サンプリング周波数は 5[Hz] とし，発声者や状況によるばらつきを豊富に含むデータを用いる．HMM の構造は，状態数を語音数と同じ 6 状態，接続は全結合とする．識別を行う際には，あらかじめ推定した HMM のパラメータを用いてゆう度の計算を行う．識別対象の単語数は 20 語と定めた（表 2 参照）これらは，使用者が相手と意思疎通を行う際に最低限必要となる単語（“はい”，

表 2 開発したシステムで使用可能な単語
Table 2 20 words used in the developed system.

おはよう	こんにちは	こんばんは	ありがとう
さようなら	たべもの	はい	いいえ
うれしい	のみもの	いただきます	げんき
いたむ	おねがい	トイレ	すき
きらい	ねむたい	つかれる	おかあさん

“いいえ” など) や、欲求を表現するために必要な単語 (“トイレ”, “のみもの” など) である。

3.4 音声生成処理

音声生成処理には、音声合成ソフトウェア (IBM (株) ProTALKER 97)[16] を利用した。このソフトウェアはパソコン上で動作し、自然な音声で日本語の文章を読み上げることができる。また、男性・女性、読み上げる速さ、声の高さ (ピッチ周波数)、声の大きさなどを自由に設定することが可能である。

4. 実 験

提案手法の有効性を検証するために実験を行った。被験者は、健康な男子大学生 4 名 (A ~ D, 22 ~ 24 歳) と、言語障害を有する喉頭切除者 1 名 (E, 45 歳) である。喉頭切除者は頸部に人工呼吸器を装着しており、従来式の電気式人工喉頭 [3] は使用することができない。EMG 信号は、図 3 に示す $L = 5$ CH から計測した。それぞれ 1CH: 咬筋, 2CH: 顎二筋, 3CH: 下唇下制筋, 4CH: 大頬骨筋, 5CH: 口角下制筋である。これらは、実験前に被験者に自然な状態で発声を行わせ、電極の装着が比較的容易で EMG 信号の振幅が顕著に現れた部位である。

実験では、まず LLGMN による語音識別の精度について調べ、次に HMM を用いて単語識別を行った。結果を以下に示す。

4.1 語音識別実験

LLGMN を用いた語音識別結果の一例を図 4 に示す。この図は喉頭切除者 (被験者 E) の結果で、上から順に、5 チャンネル分の EMG 信号、EMG 信号から求めた推定筋力、エントロピー、語音識別結果を示す。影の部分は、推定筋力 $F_{EMG}(t)$ がしきい値 F_d を越えないために無発声と識別されたことを示している。図より、一部のチャンネルでは、EMG 信号の値が小さい傾向が見られる。この患者は、日ごろ柔らかな食物を摂食しており、咬筋などの顎関節に関与する筋が減少しているためと考えられる。「い」の発声の開始と終了、「う」の発声の前半に誤識別が見られるが、全体では 87.62% の識別精度が得られた。

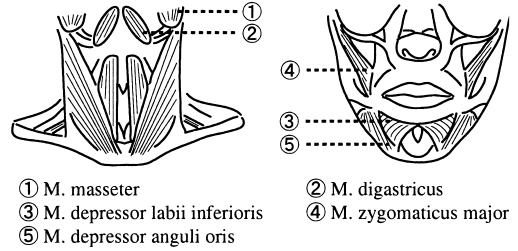
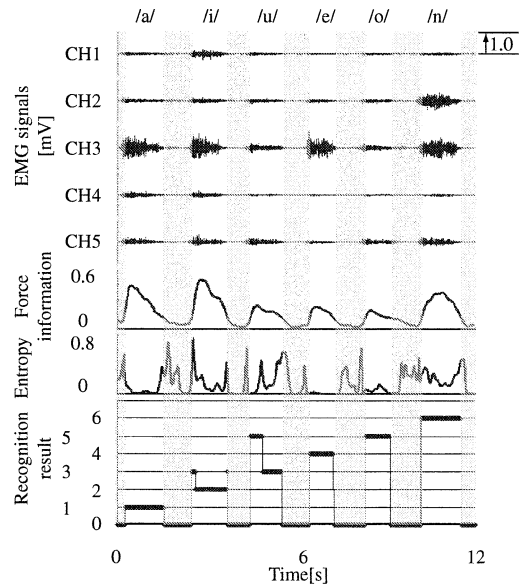


図 3 EMG 信号の計測部位
Fig. 3 Location of electrodes.



Result : 0. no sounds, 1./a/, 2./i/, 3./u/, 4./e/, 5./o/, 6./n/

図 4 喉頭切除者による語音識別結果
Fig. 4 Examples of the phoneme recognition result with the laryngectomee.

表 3 は、自然発声における語音ごとの識別結果をまとめたものである。被験者は A (健康者) で、このときに初めて語音識別を試みた。実験前に練習などは特に行っていない。この場合、「う」と「お」、「い」と「え」の語音間で誤識別率が高いことが分かる。この傾向は他の被験者も同様であった。これらの語音を発声する際に口の形状が類似していることから、EMG 信号のパターンが類似していたものと考えられる。これらの誤識別への対策としては、使用者が語音発声時にやや誇張した発声を意識することが効果的である。以降の実験では、それぞれの語音について、「あ」: 口をやや大きめに開く、「い」: 軽く歯をかみ合わせる、「う」: 口を少しすぼめ気味にする、「え」: 通常発声、「お」:

表 3 自然な発声における語音識別結果

Table 3 Recognition results of a normal vocalization.

		Recognition results					
		/ a /	/ i /	/ u /	/ e /	/ o /	/ n /
Phoneme	/ a /	70.63	0	2.88	0	26.49	0
	/ i /	0	53.24	0	38.51	0	8.25
	/ u /	0	0	39.18	0	51.27	9.55
	/ e /	0	33.72	7.86	46.55	11.87	0
	/ o /	4.5	0	27.16	0	68.34	0
	/ n /	0	0	13.2	0	0	86.8

あごをやや下げ気味にする,「ん」: 歯を少し強めにかみ合わせるなどを意識した. この際の EMG 信号のパターン変化を図 5 に示す. この図は, 特徴パターンをレーダチャート化したもので, 破線が自然発声, 実線が特徴づけた発声を示している. 被験者 A の結果であるが, 全被験者ともに似たような傾向が見られた. 自然発声時は, 「い」「う」「え」のパターンが類似しているが, 「特徴づけ」によりこの問題が解消されている. この際, 全般的に修正の規模は小さく, 「特徴づけ」がそれほど不自然でないことが分かる. ただし, 表情筋や顎部の筋活動が十分に行えない障害者の場合には, これらの「特徴づけ」が困難となり, 自然発声とは若干異なる発声動作が必要なものもある. しかしながら, 使用者が必要度に応じてそのことを許容できれば, 本装置を有効に使用できるものと考えられる.

表 4 は, このような方法で語音識別を行った結果である (被験者 A). 表 3 と比較して識別精度が著しく向上していることが分かる.

表 5 は, 被験者 5 名による実験結果をまとめたものである. 各値は, 5 試行の平均, 標準偏差を示している. 健常者 4 名の結果は, 各被験者ともに 90% 以上の識別率を実現しており, 母音及び撥音を高精度に識別することが可能なことが分かる. 喉頭切除者の結果は, 撥音の場合は 100% に近い安定した識別結果が得られたものの, 母音に関しては, 82~92% にとどまった. ただし, この被験者の場合, 実験初日には, 3 語音しか識別できなかったものが, 実験を重ねるにつれて 4, 5, 6 語音と増加してきた経緯があり, リハビリ訓練などを重ねることにより識別精度を向上できる可能性がある.

4.2 単語識別実験

次に, HMM を用いた単語識別の一例を図 6, 図 7 に示す. まず, 図 6(a), (b) は, 1 秒程度の間隔を空けながら語音発声を行ったもので, (a) は被験者 B, (b)

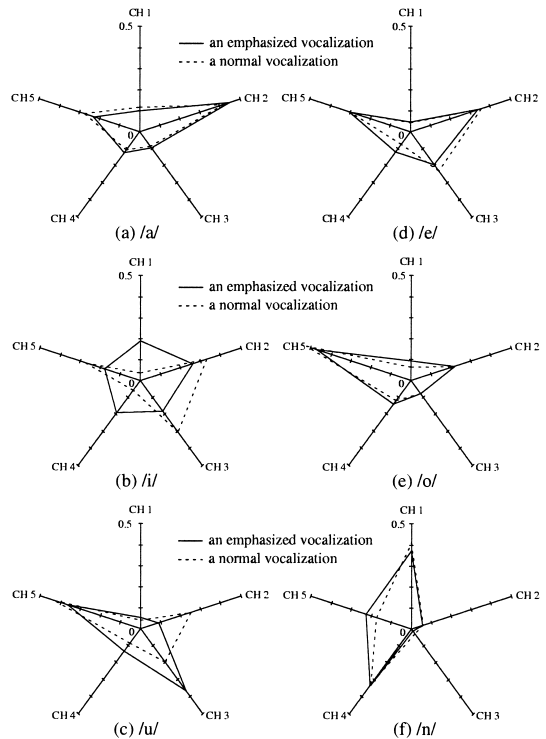


図 5 特徴づけた発声による EMG パターンの変化
Fig. 5 Changes of the EMG pattern based on an emphasized vocalization.

表 4 特徴づけた発声における語音識別結果

Table 4 Recognition results of an emphasized vocalization.

		Recognition results					
		/a/	/i/	/u/	/e/	/o/	/n/
Phoneme	/a/	97.51	0	0	0	2.49	0
	/i/	0	96.67	0	3.33	0	0
	/u/	0	0	95.29	0	4.71	0
	/e/	0	5.03	1.86	93.01	0	0
	/o/	0	0	5.18	0	94.82	0
	/n/	0	0	0	0	0	100

表 5 5 名の被験者による語音識別結果

Table 5 Recognition results for five subjects.

	/a/	/i/	/u/	/e/	/o/	/n/	Ave.	SD.
Subject A	97.51	96.67	95.39	93.01	94.82	100	95.75	3.01
Subject B	98.62	94.81	92.8	95.17	97.54	100	96.49	2.68
Subject C	100	97.94	95.33	97.29	100	100	98.42	1.92
Subject D	92.44	89.84	93.19	92.36	91.21	97.58	92.77	2.63
Subject E	91.24	83.73	85.65	82.47	84.24	98.39	87.62	6.01

Mn. : Mean SD. : Standard Deviation

は被験者 E の結果である．識別した単語は，(a) “ありがとう”，(b) “うれしい”，である．このように 1 語音ずつ順番に発声する方法は，連続的に発声するよりも識別精度が高くなる場合が多い．使用者の障害度な

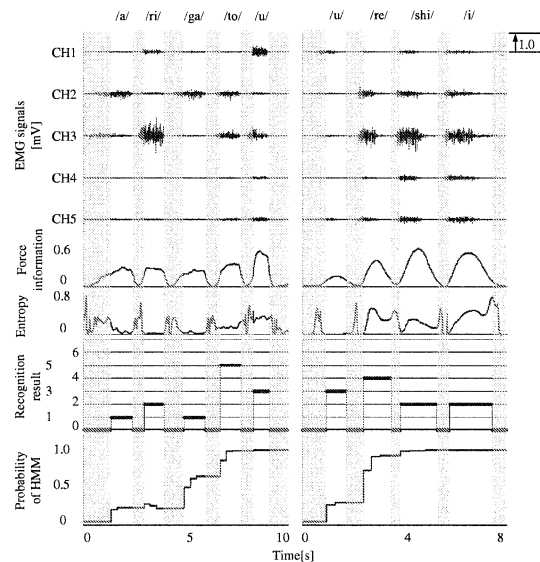


図 6 単語識別結果

Fig. 6 Example of a word recognition result.

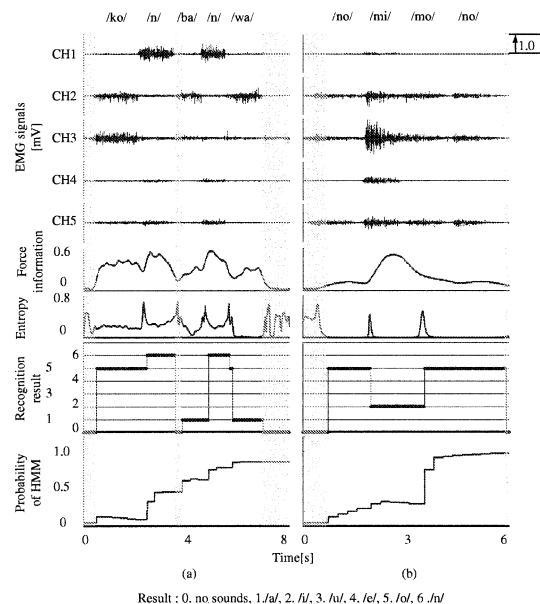


図 7 自然発声における単語識別結果

Fig. 7 Example of a word recognition result with normal vocalization.

どが重い場合には，有効な方法である．

いずれの図も，上から順に 5 チャンネル分の EMG 信号，EMG 信号からの推定筋力，エントロピー，語音識別結果，HMM の確率推移（対象とする単語について）を表している．HMM による事後確率の計算は，5[Hz] のタイミングで実施した．(a) の場合，“ありがとう”の“あり”までを発声した段階では，登録単語中の“はい”の母音列と競合するために HMM の確率が低くなっているが，順に入力していくにつれて，HMM の確率は大きくなり正しい単語識別が実現されている．(b) の被験者 E の場合も，LLGMN の識別精度は安定しており，EMG 信号から単語識別が可能なが分かる．

図 7 は，連続して発声を行った際の識別結果の一例である．(a) は被験者 A，(b) は被験者 E の結果で，発声した単語は，(a) “こんにちは”，(b) “のみもの”である．(a) では，語音識別の段階において多少の誤識別が見られたが，HMM のロバスト性により単語を正しく識別している．(b) の結果に誤識別はなく，喉頭切除者にとっても本システムが利用可能なが分かる．

また，システムに組み込まれた全単語に対する識別結果を表 6 に示す．被験者は，A～D で各単語とも 5 試行を行いその正解率を記載している．この結果から，各被験者ともに高い識別精度が実現できていることが確認できる．いくつかの誤識別が見られるが，それが

表 6 4 人の被験者による単語識別結果
Table 6 Word recognition results with four subjects.

	Subject A	Subject B	Subject C	Subject D
/ohayou/	100	100	80	100
/konnichiwa/	100	80	100	80
/konnbannwa/	100	100	100	100
/arigatou/	100	100	80	100
/hai/	80	80	100	80
/iie/	100	100	100	80
/suki/	100	100	80	80
/kirai/	100	100	100	100
/sayounara/	100	80	100	100
/toire/	100	100	100	100
/tabemono/	100	100	100	100
/nomimono/	100	100	100	100
/gennki/	100	100	80	100
/tsukareru/	100	80	100	100
/ureshii/	100	100	100	80
/itadakimasu/	80	100	60	100
/itamu/	100	60	100	80
/onegai/	100	100	80	60
/nemutai/	100	100	100	100
/okaasan/	100	100	60	100
total	98	94	91	92

一つの単語に集中することはなく、全単語ともに発声が可能なが分かる。

5. む す び

本研究では、発声機能障害者のための新しい意思伝達装置の開発を目的とし、EMG 信号を用いた新しい代用発声法を提案した。健康な男子大学生及び喉頭切除者の計 5 名による実験を行った結果、精度良く語音・単語を識別できることが示された。

今後は、EMG 信号からのピッチ情報の抽出を試みるとともに、発声できる単語数の増加を検討している。この際、同一構造をもつ語音列が考えられるが、希望する単語を選択するためのインタフェース画面などを用意し、操作性の向上を図ることを考えている。なお、本研究の一部は NEDO 産業技術研究助成事業 (01A17001b) の助成を受けて実施したものでここに改めて謝意を表します。

文 献

- [1] 厚生労働省社会・援護局障害保健福祉部企画課,
<http://www.mhlw.go.jp/index.html>
- [2] 高橋次夫, “我が国無喉頭音声の変遷と現状—銀鈴会を中心として,” 音声言語医学, vol.24, pp.184–189, 1983.
- [3] 関 恵貞, 高橋 誠, 西澤典子, 西澤伸志, 伊福部達, 犬山征夫, “胸骨舌骨筋の筋電位による電気式人工喉頭の制御,” ME 誌, vol.32, no.4, pp.297–305, 1994.
- [4] 橋場参生, 上見憲弘, 及川雅俊, 山口悦範, 須貝保徳, 伊福部達, “抑揚制御機能を備えた電気式人工喉頭の製品化と評価,” 信学論 (D-II), vol.J84-D-II, no.6, pp.1240–1247, June 2001.
- [5] E.M. Mullar, “Strain gauge transduction of lips and jaw motion in the midsagittal plane: Refinement of a prototype system,” J. Acoust. Soc. Am., vol.65, no.2, pp.481–486, 1979.
- [6] R.L. Goode, “Artificial laryngeal devices in post-laryngectomy rehabilitation,” The Laryngoscope, vol.85, pp.677–689, 1975.
- [7] 村木 茂, 角田耕一, 杉江 昇, “代用発声のための子音判別法に関する基礎的研究,” 信学技報, MBE83-108, 1984.
- [8] 角田耕一, 杉江 昇, “音声合成方式発声代行システム—筋電位信号からの母音の判別と発声,” 電学論 (C), vol.105-C, no.2, pp.25–32, 1985.
- [9] 真鍋宏幸, 平岩 明, 杉村利明, “筋電信号を用いた無発声音声認識一定常状態における母音の識別,” 情報処理学会シンポジウムシリーズ, vol.2002, no.7, pp.181–182, 2002.
- [10] 辻 敏夫, 伊藤宏司, 長町三生, “義手制御を目的とした多チャンネル EMG 動作識別法,” 信学論 (D-II), vol.J70-D-II, no.1, pp.207–215, Jan. 1987.
- [11] 福田 修, 辻 敏夫, 金子 真, “EMG 信号を利用した手動制御型人間支援マニピュレータ,” 日本ロボット学会誌, vol.18, no.3, pp.387–394, 2000.
- [12] 辻 敏夫, 重吉宏樹, 福田 修, 金子 真, “EMG 信号に基づく前腕動力義手のバイオミメティック制御,” 日本機械学会論文集 (C), vol.66, no.648, pp.2764–2771, 2000.
- [13] T. Tsuji, O. Fukuda, H. Ichinobe, and M. Kaneko, “A log-linearized Gaussian mixture network and its application to EEG pattern classification,” IEEE Trans. Syst. Man Cybern. C, Appl. Rev., vol.29, no.1, pp.60–72, 1999.
- [14] L.R. Rabiner, “A tutorial on hidden Markov model and selected application in speech recognition,” Proc. IEEE, vol.77, no.2, pp.257–286, 1989.
- [15] 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.
- [16] 日本 IBM 株式会社, <http://www.ibm.com/jp/>
(平成 15 年 8 月 29 日受付, 16 年 3 月 3 日再受付)



福田 修 (正員)



藤田 真治



辻 敏夫 (正員)

2000 広島大学大学院工学研究科博士課程後期了。1997～1999 の期間、日本学術振興会特別研究員 (DC1)。2000 通商産業省工業技術院機械技術研究所入所, 2001 独立行政法人産業技術総合研究所へ転任, 現在に至る。博士 (工学)。ニューラルネット、ヒューマンインタフェース、電動動力義手などの研究に従事。日本機械学会, 日本人間工学会, 日本ロボット学会等各会員。

2001 岡山大・工・システム卒。現在、広島大学大学院工学研究科博士課程前期課程在学中。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。日本機械学会の学生会員。

1985 広島大学大学院工学研究科博士課程前期了。同年広島大学工学部助手。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。工博。計測自動制御学会学術奨励賞 (1986), 論文賞 (2002), バイオメカニズム学会論文賞 (1990), 日本義肢装具学会論文賞 (2000), 日本医科器械学会論文賞 (2003) などを受賞。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。