

4 E - 6

# 音声認識エンジンを利用した WWWのフォームへの音声入力アクセス

遠山純司<sup>†</sup> 草薙信照<sup>‡</sup> 太田忠一<sup>†</sup>

大阪国際大学大学院経営情報学研究科<sup>†</sup> 大阪経済大学経営情報学部<sup>‡</sup>

## 1 はじめに

WWW(World Wide Web)の普及により手軽にマルチメディアデータベースにアクセスできるようになってきた[1]。このWWWにおいてフォームはユーザの登録や情報検索のためのキーワード入力等、対話的な処理に広く利用され、重要な役割を演じている。

一方、音声は日常的コミュニケーションを行うための重要な手段の一つである。この音声コンピュータの入力手段として用いることによって誰にでも容易に、自然な入力が可能となる[2][3]。最近のコンピュータの高性能化と音声認識技術の向上により一般ユーザでも身近に音声入力が利用できるようになっている[4]。

ここでは、WWWのフォームにおける音声入力アクセスについて検討し、それに基づき音声認識エンジンを用いて作成した音声入力ブラウザの概要について報告する。

## 2 WWWのフォーム

WWWはテキストや画像などのマルチメディア情報をリンクで結びつけたハイパーメディアシステムである。WWW上の情報はHTML(Hyper Text Markup Language)によって記述されており、WWWブラウザはWWW上のHTML文書を表示するためのソフトウェアである。WWWのフォームもHTMLを使って表現され、例えばテキストボックスを表示させるためには、

表1 HTML3.2におけるフォームタグ [5]

タグ	TYPE	内容
INPUT	text	一行だけのテキスト入力
	checkbox	複数選択が可能なチェックボックス
	radio	複数項目から一つの項目を選択
	submit	データをサーバーに送る実行ボタン
	reset	データをすべてクリアするボタンの作成
SELECT	—	選択メニューの作成
TEXTAREA	—	複数行に渡ってテキストを入力

`<input type="text" name="name">`

のように記述する必要がある。フォーム作成の際によく用いられるHTMLタグを表1に示す。

## 3 フォームにおける音声アクセス

### 3.1 コマンドとテキスト入力の切替

音声でフォームにアクセスする場合、WWWブラウザ内の項目移動等のコマンドとテキストボックス等へのテキスト入力を区別する必要がある。制御や操作のコマンドとテキスト入力の切替え方法には次のような方法が考えられる。

- (1) 特定の単語の入力
- (2) 一定時間の無音の検出
- (3) マウスなど他の入力方法の使用
- (4) 手を叩くなど音の特性の著しい変化の検出

### 3.2 音声アクセス方法

現在WWWブラウザの入力手段としてキーボードとマウスが主に用いられているが、音声を用いることにより、ユーザはより自然な対話が可能となる。フォームにおいてキーボードは主にテキスト入力と項目を順番に指定する順次指定に用いられ、マウスは項目の直接指定やチェックボックスなどの選択に用いられている。音声によるフォームへのアクセス方法として、次のような方法が考えられる。

#### (1) 順次アクセス

もっとも単純なアクセス方法で、「次に」等の発声によってフォームの入力項目を順次移動する。キーボードによるタブキーと同じような選択方法である。フォームの項目選択に使われる音声コマンドがごく少数であるため認識率が高くなるが、項目選択に手間がかかる。

#### (2) 見出しアクセス

フォーム項目の見出しを発声して、その項目を選択する。項目を直接的に選択できるが、見出しのない項目では選択が難しい。順次アクセスや他の入力手段と組み合わせる必要がある。

Spoken Language Access to the Form on WWW utilizing Voice Recognition Engine

Junji TOHYAMA<sup>†</sup>, Nobuteru KUSANAGI<sup>‡</sup>, Chuichi OTA<sup>†</sup>

Osaka International University<sup>†</sup>

Osaka University of Economics<sup>‡</sup>

表2 フォームへの音声アクセス方法

アクセス方法	利 点	欠 点
順次アクセス	-操作が単純 -使われるコマンド数が少なく済む	-項目の直接選択が不可能 -特定の選択に時間を要す
見出しアクセス	-項目の直接選択可能 -操作が直感的で自然	-見出しのないフォームでは選択不可能
文章アクセス	-コンピュータとの自然な対話が可能 -見出し入力に比べ操作がさらに自然	-高い文章理解能力が必要 -画面上での入力確認が困難

### (3) 文章アクセス

人間と会話しているかのようにコンピュータに自然語で発声し、コンピュータが文章を解析して各項目にテキスト文を入力する方法。コンピュータと自然な対話が可能であるが、自然語を解析するには人工知能などの高い解析能力が要求される。

上述のアクセス方法の利点と欠点を表2に示す。

## 4 音声アクセスの可能なブラウザ

フォームでの音声アクセス方法の検討を基に、音声アクセスが可能なブラウザを作成した。今回はフォームでの音声アクセスに焦点を絞っているため、ブラウザとしての基本的な動作にはMicrosoftのWWWブラウザコントロールを用いることとし、音声アクセスのためにフォーム専用のサブウィンドウを別に作成することとした。

サブウィンドウは図1に示すように、「フォーム入力開始」と発声することで開き、「フォーム入力終了」で終了するようにした。

基本的なアクセスは順次アクセスとし、「次に」、「戻る」等の発声によりカーソルがフォーム上を移動するようにした。見出しアクセスについても、「見出し」と発声した後、各項目の見出しを発声することにより見出しアクセスができる。

「テキスト入力開始」という発声でテキストボックスへのテキスト入力開始され、「テキスト入力終了」という発声によって終了する。また、ラジオボタン等の選択には「選択」と発声する。

すべてのフォーム入力が終了したときには「フォーム入力終了」でサブウィンドウを終了させ、メインウィンドウに戻る。

なお、音声認識エンジンにIBM ViaVoice Developer Tools V4.3Jを使用した。

## 5 おわりに

本論文では、WWWにおいて重要な役割を担っているフォームに焦点を絞り、音声によるフォームへの音声入

力アクセス方法を検討し、これを基に音声アクセスが可能なWWWブラウザの実現を図った。今後、完成度を高めるとともに音声合成システムなどを付加し、よりユーザーに自然なインタフェースをもつシステムの開発を進めて行きたい。

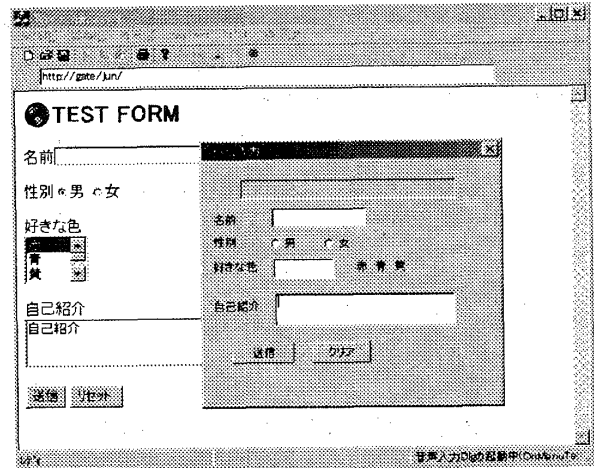


図1 開発した画面例

## 参考文献

- [1] <http://www.nw.com/>
- [2] 新田恒雄"マルチメディア・サービスとマルチモーダルUI", 音声言語処理 1-5, pp.31-38, (1994-05)
- [3] 田村 博編"ヒューマンインタフェース", オーム社, (1998)
- [4] <http://www.ibm.co.jp/voiceland/>
- [5] [http://www.abacuss.com/HTML\\_help/form.html](http://www.abacuss.com/HTML_help/form.html)