

高齢者向け情報入出力装置の構築

Construction of an interactive device for the elderly people

三隅隆也 北山一郎

MISUMI Takaya, KITAYAMA Ichiro

キーワード：

コミュニケーション、タッチパネル、マンマシンインターフェース、音声認識

Keywords:

Communication, Touch panel,

Man-machine interface, Speech-recognition

Abstract:

Multi-media information apparatus such as a personal computer, a phone with TV and training machine with excise-function are spreading in hospitals, facilities and each home, but the handling of operational devices such as a keyboard of these apparatus, display, 10-key feels resistance to a big burden or apparatus itself for elderly people and disable people and seems unwilling to operate.

We have developed the system which can communicate various desired intentions using a touch panel sensitively and visually.

In this annual we have developed several methods of a voice recognition. We hope to make a life support spread by using these systems.

1 はじめに

パソコン(PC)やテレビ電話などのマルチメディア情報機器および運動機能訓練機器が施設や各家庭で普及しつつあるが、これらの機器のキーボード、ディスプレイ、テンキーなどの操作機器の取り扱いが高齢者等に対し大きな負担となっている。この問題を解決するために、前年度までに「高齢者・障害者のための双方向型入出力装置」を開発した。この入出力装置は、PCにタッチパネル、スキャナ、CCDカメラを組み合わせ、タッチパネルに触れることによってキーボードやマウスを使うことなく、スキャナから文字や写真や創作した絵を入力で

き、さらにCCDカメラから本人の顔写真などが入力できる。その結果、タッチパネルを使うことで入力が容易になり、かつ入力した文字・絵・写真が外部へ送信ができるため高齢者等が家族との会話・高齢者同士の会話・コミュニティセンターでのふれ合いの場等ですぐにでも使いたいとの要望が高かった。一方でタッチパネルに触れたことを音声で確認すること、タッチパネルに触れるのではなく音声で操作することへの要望も高かった。このため、今年度は、音声認識や音声合成の機能を付加し操作性を拡張し・取扱いの容易性を考慮し普及可能な装置を構築することを目的とした。

2 研究の概要

本開発においては、前年度の「双方向型入出力装置」の機能拡充を図り以下の機能が出来ることを目指した。(1)マウス・キーボードを使わずタッチパネルの画面に触れるだけで操作できること、(2)写真、創作絵画、映像(ライブ静止画))など画像を中心とした電子メールをやりとりすることにより、高齢者等に親しみやすい情報端末を提供すること、(3)タッチパネルの操作が不自由な人でも、音声により操作ができること、(4)音声合成による各種メッセージの発声により操作の確認ができること。

2.1 ハードウェア構成

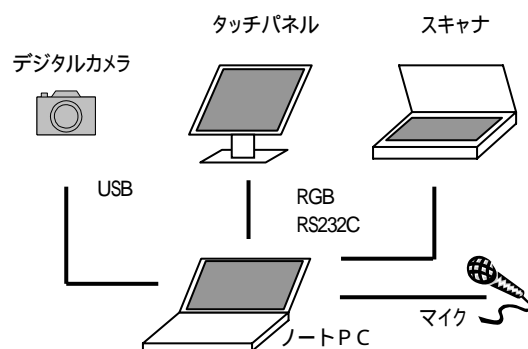


図1 1 システム構成

Fig. 1-1 System configuration



図1 - 2 システム構成
Fig.1-2 System configuration

本システムのハードウェア構成を図1に示す。

- PC(パーソナルコンピューターノートブック型)
- タッチ・センサー付き液晶カラーモニター15 インチ
- スキャナ 光学解像度 1200dpi × 2400dpi
- デジタルカメラ 記録画素数 320 × 240 pixels

2.2 ソフトウェア構成

本システムでは、音声認識するためのソフトと従来のタッチパネルによる画面操作を行うためのソフト開発を行った。下記に開発ソフトを記載するが、本稿では音声認識の開発ソフトを重点的に述べる。

[開発ソフトウェアの内容]

1) 音声認識機能

NEC 製 Voice SDK V4.6(以下 SDK)の活用して
母音認識アルゴリズムの開発する

2) 画面操作支援機能

従来の機能に音声による確認・警告・指示を行なう機能を付加する

- 構内メーリング・システム
- 構外メーリングの実現
- 書庫ハンドラー
- 掲示板およびFAXブラウザー
- リハビリテーション・ゲーム

3 音声認識

今回、音声認識には、以下のような方式を採った。

- 1) 前述のように語句(アイコン)を読み上げて画面操作を行う方式
- 2) 上下左右(うえ・した・みぎ・ひだり)を発生しカーソルを動かしてカーソルが所定のアイコンの位置にきたとき画面操作を行う方式

- 3) 母音(あ・い・う・え・お)を発生しカーソルを動かしてカーソルが所定のアイコンの位置にきたとき画面操作を行う方式

3.1 音声認識 1

前述、音声認識 1)、2)の方式は音声認識エンジン(NEC 製 Voice SDK)を基盤として以下の機能を実現した。

- ・ アプリケーションのコマンド・ボタンの読み上げによる実行
- ・ カーソルの誘導とクリック動作の音声誘導

ソフトウェア処理は、マイク入力を介して取り込んだ音声信号を PC で解析し、音声認識エンジンを用いて登録語彙を識別し、それぞれの識別結果により画面上のアイコンによるコマンドの操作および音声によってマウス・カーソルポインターの誘導により画面の操作を行なう。

市販の大多数のソフトウェアでは、マウスを用いてカーソルポインターを移動し、ボタンを押す(クリック)ことで各アプリケーションにメッセージを伝達する方式を採用しているため、本プログラムをマウスの代用とすることができ、タッチパネルを利用できないユーザーに対して、音声による入力手段を提供することが可能となる。

ここでは、前半に画面操作を音声で行うための仕組みについて画面を見ながら説明し、後半で音声認識の方法について述べる。

3.1.1 音声認識による画面操作

システムを立ち上げると図2の画面が表示される。従来は、この画面に触れてログインを行っていたが、これを左肩の数字「いちばん」、「にばん」などと音声が発生してログインを行うことができる。

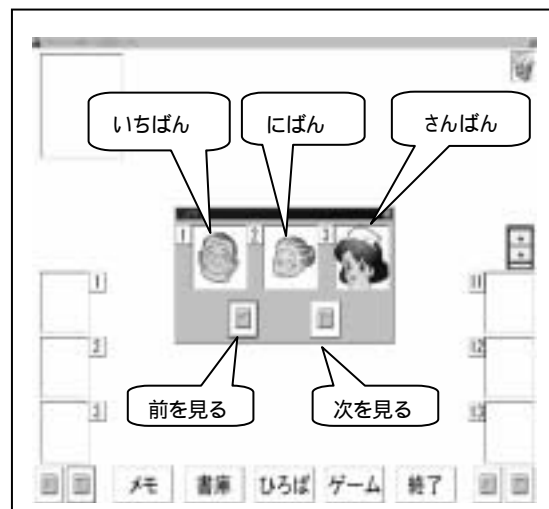


図2 ログイン画面
Fig.2 Screen of login

次に、主画面(図3)が表示されると、メールアイコンや送りボタン、宛先ユーザアイコンなどに番号が付くので、その番号を読み上げることで操作を行うことができる。「メモ」、「書庫」、「ひろば」などは、そのまま「めも」、「しょこ」、「ひろば」と読み上げることで操作できる。

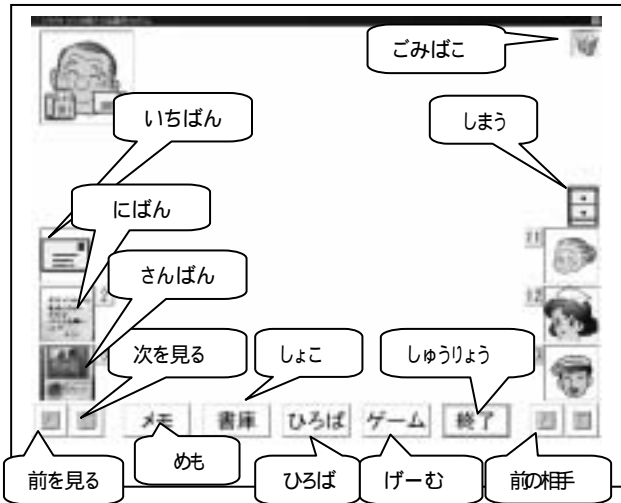


図3 音声入力の主画面

Fig.3 Main Screen of Voice Recognition

3.1.2 語句の登録

語句の登録は、図4の「新規辞書」の登録機能を用いて図2、図3にあるような語句を登録していく。今回はアイコンの数47個の語句が登録されている。



図4 語句の登録

Fig.4 Registration of the words and phrases

3.1.3 音声によるカーソル移動

図5に示すように「うえ・した・みぎ・ひだり」を発生しカーソルを主画面上を動かし所定のアイコンの位置にきたら「OK」を発生することにより画面をクリックすることができる。そのときの語句の登録(図6)は、前述のように辞書登録機能「新規辞書」を用いて登録することができる。

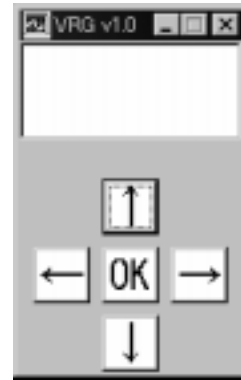


図5 音声によるカーソル移動

Fig.5 Cursor movement by a sound

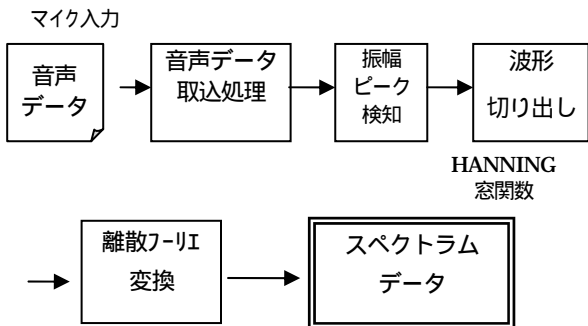


図6 語句の登録

Fig.6 Registration of the words and phrases

3.2 音声認識2

母音(あ・い・う・え・お)を発生しカーソルを動かしカーソルが所定のアイコンの位置にきたとき画面操作を行う方式(母音認識)は、マイク入力した音声以下を以下の処理方式を採って処理している。



また、カーソル誘導処理は以下の処理を行っている。

3.2.1 母音の教示

教示する母音を選択し、母音を発声する。図7は「あ」の例である。

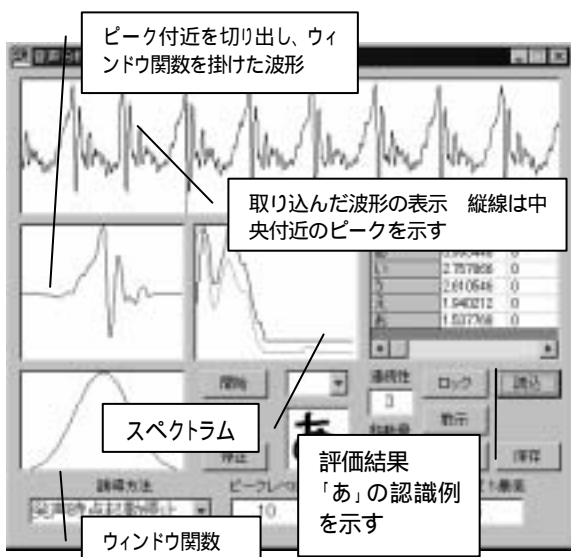


図7 母音の教示

Fig.7 Instruction of a vowel sound

図7に示す母音の教示は以下の方法で行っている。

- 1) 音声信号を100ミリ秒のサンプリング間隔で取り込む
 - 2) 波形の表示区間中央付近でピークを検出する(図の縦線部)
 - 3) ピークを中央としウィンドウ関数を掛けて波形を切り出す
 - 4) 切り出した波形を高速フーリエ変換しスペクトラムを求める
 - 5) 基本調波の40倍までを選択し、母音特徴量(音声パラメータ)として登録する
- 上記のように、あらかじめユーザの母音が登録さ

れ「あ」が発声されると図8のように認識される。

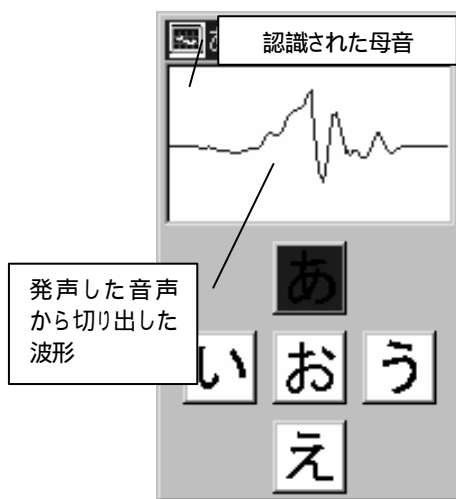


図8 母音の認識

Fig.8 Recognition of a vowel sound

認識した母音に対して、次の処理を行なう。

- 【あ】カーソル・ポインタを上に移動
- 【い】カーソル・ポインタを左に移動
- 【う】カーソル・ポインタを右に移動
- 【え】カーソル・ポインタを下に移動
- 【お】所定のアイコン上でマウスの左ボタンを押下しクリックした状態になる

4 音声合成の仕組み

NEC製の音声合成システム(Voice SDK V4.6)を組み込み、これとのインターフェースを行うプログラム(Speech Board)を開発した。クリップ・ボード(WINDOWSのメッセージ交換機能)経由で他のアプリケーションから任意の「文字列」を受信することにより文字列を合成音声で発声する。

高齢者向け情報入出力装置用ソフトウェア(STERM V2.0[®])には、Speech Boardへのメッセージ送信機能のみを組み込んでおり、STERM起動時にSpeech Boardも同時に起動し、音声合成を利用できる状態にする。図9に、その例を示す。

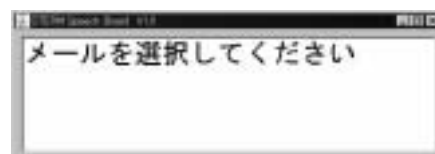


図9 音声合成の例

Fig.9 An example of a voice synthesis

例えば、以下のメッセージが発声される。

- ・送付先ユーザアイコン：上の（中央の、下の）ユーザに送ります（画面表示）
- ・ごみ箱、書庫、デスクアイコン：～へ移動します（画面表示）
- ・確認画面など：画面上のメッセージ

5 評価試験

今年度は、以下の方式による音声認識の研究開発を中心に行い、これらの方式の評価試験を実施した。（図 10）

5.1 評価方法と認識結果

- ・対象者：30 歳代の男性（頸椎損傷者）
- ・評価方法
 - 1)画面のアイコンの語句を読み上げて画面の各操作を行う方式
 - 2)カーソルを上下左右（うえ・した・ひだり・みぎ）と発生して動かしたカーソルが所定のアイコンの位置にきたとき画面の各操作を行う方式
 - 3)母音（あ・い・う・え・お）を発生しカーソルを動かしてカーソルが所定のアイコンの位置にきたとき画面の各操作を行う方式
- ・認識結果

方式 1)：通常の音声の発生で認識率は、ほぼ 100% に近い。

方式 2)：通常の音声の発生で認識率は、ほぼ 100% に近い。しかし、画面上のカーソルの移動時間とカーソルが所定のアイコンの位置にきたときに「OK」と発生し画面をヒットさせるのに、移動時間を 2 段階にしているが、早すぎると「OK」を発生しヒットするのに追従できなくなり、遅すぎると話者がいらだちを感じるので話者の特性に併せた工夫が必要である。ただ、学習効果が出てくるので操作も容易になる。

方式 3)：あらかじめ話者が登録した母音「あ・い・う・え・お」の単音と発生した音声特徴とが一致した場合にカーソルの移動・画面のヒットを行っているが、音声の特徴を十分に捉えておらず認識率が低い。

5.2 対象者の意見

評価願った結果、以下の意見があった。

- 方式 1)および方式 2)については、認識率が高く、いままで使ったことのある音声認識装置に比べ非常に優れており十分に使える。
- ・母音の発生の方は、認識率が低いので使い難いが、語句の発生ができにくい障害者には、一語発生で

- 使えるような工夫をしたら良いと思う。
- ・音声であると姿勢を固定しなくても操作できるので便利である。
- ・本装置は、タッチパネルでも音声認識でも使えるので NEDO の支援で開発されている環境制御装置にも適用できる。
- ・音声で文章が入力できれば非常に良い。



図 10 評価風景

Fig.10 Scenery of Evaluation

6 あとがき

3 通りの方式で音声認識を行った。音声認識エンジン（NEC 製 Voice SDK）をもとに実現した下記機能、

- ・画面のコマンド・ボタンの読み上げによる実行
- ・マウス・カーソルの誘導とクリック動作の音声誘導

は、今回は約 50 語の音声認識であったが、ほぼ 100%に近い認識率であり、100 語ぐらいは高い認識率を得られると想われるため音声による有効な入力手段を提供することが可能と考える。このため、本方式による音声認識の活用を図りたい。

参考文献

- 1) 小林他：介護ロボットのための音声対話システムの検討 音講論秋季 pp59-60(2001.10)
- 2) 音声認識技術の情報家電への応用 SICE SI 部門講演会 pp73-74(2002.12)
- 3) 平成 13 年度兵庫県立福祉のまちづくり工学研究所報告書 pp105～110