

音声対話を含むマルチ・モダル型ユーザ・インターフェースの 在宅健康管理システムへの適用

鳥原 信一

日本アイ・ビー・エム株式会社 東京基礎研究所

音声対話を含むマルチ・モダル型ユーザ・インターフェースを、HTML をベースとするアプリケーション・フレームワークとして実現した。このアプリケーション・フレームワークの上に在宅健康管理システムを構築した。

音声認識は、コマンド・ナビゲーション・モードを用いた。音声対話誘導により、発話内容を制限し、しかし、発話表現を豊富にカバーすることにより、音声認識率を高め、操作性と親和性を向上させている。

Speech dialog as a multi-modal user interface for Home Health Care system

Shinich Torihara

IBM Research, Tokyo Research Laboratory, IBM Japan

HTML based application framework including speech dialog as a multi-modal user interface has been developed and applied for Home Health Care (HHC) system, i.e., remote medical services between homes and medical facilities through network.

Speech dialog enables the patients, especially the aging who have not been familiar with PCs, to operate HHC terminals by themselves.

We use Command Navigation mode of ViaVoice. If proper grammar files are prepared, the ratio will be improved. We have to research on grammar file generator by dialog text corpus in order to increase the adaptability of speech dialog.

1. はじめに

音声対話を含むマルチ・モダル型ユーザ・インターフェースを、HTMLを基軸としたアプリケーション・フレームワークとして実現した。このアプリケーション・フレームワークの上に在宅健康管理システム (Home Health Care system) に適用し成功したので報告する[1]。

音声対話を、マルチ・モダル型ユーザ・インターフェースとして実現することにより、パソコン操作に困難を覚える方々、不慣れな方々も親しみやすく目的を遂行できるようになった。日々の健康管理のために、情報機器 (IT) の前に座るといふ行為に対する抵抗を少しでもなくしたいと配慮している。

2. 在宅健康管理システム

ここで紹介するのは、在宅患者と医療機関との遠隔医療システムへの音声対話を含むマルチ・モデル型ユーザ・インターフェースの適用についてである[2]。図1に本システムの全体像を示す。患者側端末と医師側端末とは公衆回線（ISDN）で接続されており、データの転送とテレビ電話を行う。患者側端末では脈拍、血圧、心電図、血中酸素濃度はオンライン接続された生体測定機器により測定され、体温は測定後入力する。測定結果は医師側端末にデータ転送される。日常数回にわたる測定スケジュールによって行うが、この中に著しい変化が自動検出されると、患者側端末と医師側端末の間でテレビ電話セッションが開かれる。テレビ電話を通して患者の状態、顔色を見ながらの処方がなされる。このシステムの端末には、タッチパネル・ディスプレイ、キーボード、マウス、ソフトウェア・キーボード、そして音声対話（音声合成、音声認識、アニメーション・キャラクタ）のユーザ・インターフェースを装備している。しかも、マルチ・モデル対応である。従って、ユーザの好み、状況によって、ユーザ・インターフェースを選択でき、使用順序、複数使用などに制限はない。このような在宅健康管理システムの当初予定している利用者は、日常生活が自立可能な方々であり、日常の健康管理が必要な方々である。つまり、毎日、患者側端末（PC）の前に数回回り、操作ができる方々である。年齢は比較的高齢者であるために、パソコン操作が初めてであったり、不慣れであったりすることから、マルチ・モデル型ユーザ・インターフェース、そして、音声対話の採用となった。音声合成、アニメーション・キャラクタ、画面表示などによる音声対話誘導により、ユーザによる発話内容を制限し、発話表現を豊富にカバーすることによって、上記のユーザにも親和性を持って操作できると思われる。本システムを構成するハードウェアおよびソフトウェアを図2に示す。

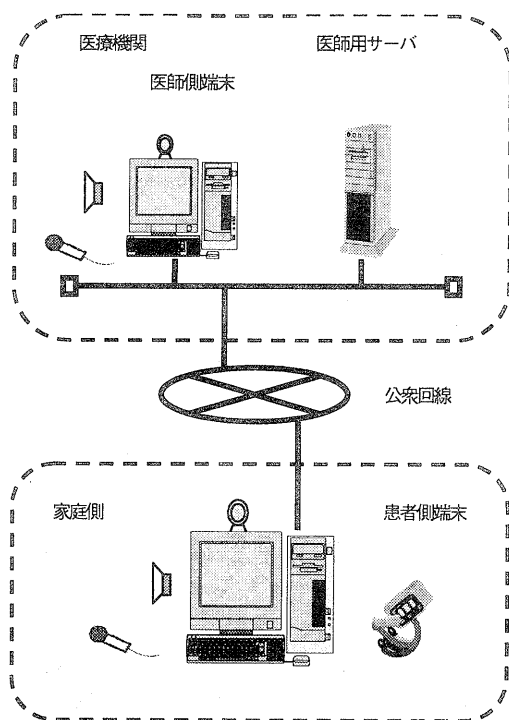


図1. システム全体像

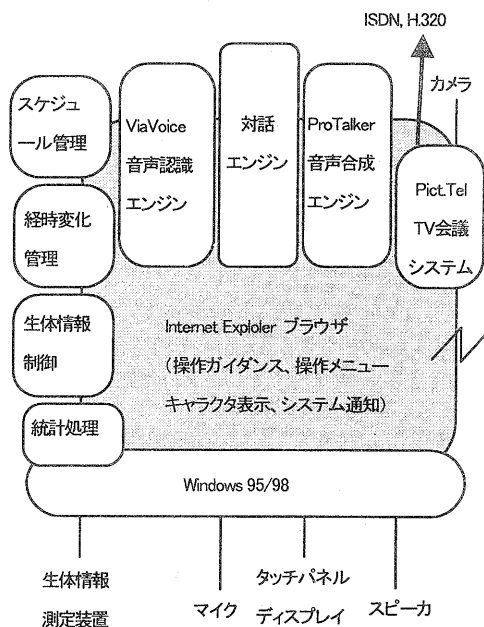


図2. ハードウェアおよびソフトウェア構成

3. 音声対話を含むマルチ・モダル型ユーザ・インターフェース

3.1 音声対話とマルチ・モダル型ユーザ・インターフェースの意義

Graphical User Interface (GUI) は、直接操作法 (direct manipulation) の一つであるように直感的に操作ができる。キーボードの未習熟な子供でも、相当操作が可能である。しかしながら、視覚に若干の障害があり、または、マウス操作に熟達できない方々にはきわめて操作が困難である。音声合成、音声認識そしてアニメーション・キャラクタによって音声対話形式で操作ができれば上記のユーザには歓迎されるユーザ・インターフェースとなる。しかも、状況によって、ユーザ・インターフェースを選択、組み合わせて使用できるマルチ・モダル型ユーザ・インターフェースは有益である。例えば、数字を入れるときには、数値キー (テン・キー) による入力の方が簡単にできる場合もある。

3.2 HTMLをベースとするアプリケーション・フレームワークの構築

HTMLをベースとしたアプリケーション・フレームワークを構築した。アプリケーション・フレームワークの概念図を図3に示す。インターネット・ブラウザを用いてHTMLによって記述されたコンテンツ (表示内容と操作の実装) を実現する。しかも各種入力デバイスによるユーザ・インターフェースをイベント制御してマルチ・モダル型ユーザ・インターフェースとしている。アプリケーション開発者はHTML、Java applet、JavaScriptの修正、作成により音声対話を含むマルチ・モダル型ユーザ・インターフェースを持ったアプリケーションが開発できる。また、ユーザも、インターネット・サーフィンの感覚で、操作可能となる。音声対話の機能を果たしている音声認識、音声合成とアニメーション・キャラクタはJavaからJNI (Java Native Interface) を通してC++言語で実装されている。音声対話制御部のイベントとデータの流れを図4に示す。

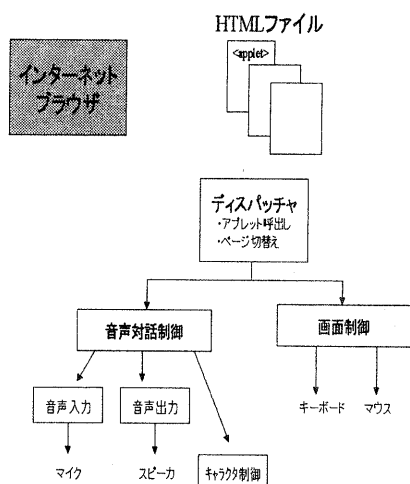


図3. アプリケーション・フレームワークの概念図

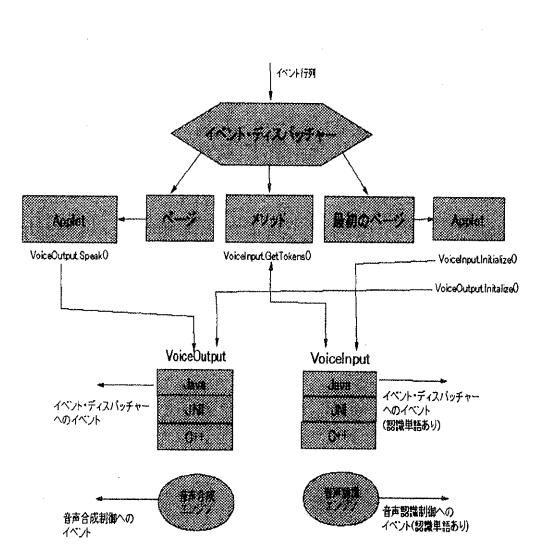


図4. 音声対話制御部のイベントとデータの流れ

3.3 コマンド・ナビゲーションによる音声認識

IBMの音声認識ソフトウェアには、ディクテーション・モードとコマンド・ナビゲーション・モードがある。本システムでは、発話される可能性のある単語、文を文法ファイルで指定するコマンド・ナビゲーション・モードを採用した。音声合成、アニメーション・キャラクタ、画面表示によるガイダンスにより、発話内容を誘導、制限し、しかし、発話表現を豊富にカバーすることにより、音声認識精度を高めることができる。文法ファイルには、認識単語、文にannotation（解釈）を付与し、これを解析して、ページ遷移などを行う。文法ファイルはシステム全体に共通なものとHTMLごとに切り替えるものがある。「体温」入力に関する文法ファイルの例を図5に示す。

```

// 001001.bnf
<root> = <taion> | <kakunin> | <center> | <modoru> | <clear>.

<taion> = <num_do> <do1?> | <num_do> <do1> <num_bu> <bu1> | <num_do> <do2> <num_bu> <bu2> | <num_bu> <bu1>.
<num_do> = <num_do31> <num_do32?> | <num_do41> <num_do42?>.
<num_do31> = "三十"; "体温度30".
<num_do32> = "五"; "体温度05" | "六"; "体温度06" | "七"; "体温度07" | "八"; "体温度08" | "九"; "体温度09".
<num_do41> = "四十"; "体温度40".
<num_do42> = "一"; "体温度01" | "二"; "体温度02".
<num_bu> = <digit0-9>.
<digit0-9> = "0"; "体温分00" | "一"; "体温分01" | "二"; "体温分02" | "三"; "体温分03" | "四"; "体温分04" | "五"; "体温分05" | "六".
<do1> = "度"; "体温".
<do2> = "点"; "体温".
<bu1> = "分"; "体温".
<bu2> = "度"; "体温".

<kakunin> = <suui i?> <yo i>.
<suui i> = <kono?> <suui i2> <do?>.
<kono> = "この".
<suui i2> = "数字".
<do> = "で".
<yo i> = "よい"; "確認" | "OK"; "確認".

<center> = <center2> <denwa?> | <center2?> <denwa>.
<center2> = <center3> <ni?>.
<center3> = "センター"; "電話" | "先生"; "電話" | "ドクター"; "電話".
<ni> = "に".
<denwa> = <TV> | <TV?> <denwa2>.
<denwa2> = <denwa3> <wo?> <call?>.
<wo> = "を".
<TV> = "テレビ"; "電話".
<denwa3> = "電話"; "電話".

```

図5. 文法ファイルの例

4. 課題と研究方針

HTMLを基軸としたアプリケーション・フレームワークを持つ、本システムを実装中にあっても、Voice XML、Voiceを含むTranscoderなど、新しい規格や提案が活発であった。重要なのは、ユニバーサル・デザインなどの観点からも音声対話ユーザ・インターフェースが評価された上で、音声対話誘導のための音声合成文の自動生成と、発話テキスト・コーパスからの文法ファイルの自動生成であると思われる。これらの研究が音声アプリケーション開発のためには、必要とされるであろう。

5. おわりに

音声対話を含むHTMLをベースとするアプリケーション・フレームワークを構築し、在宅健康管理システムに適用した。システム使用者の観察と意見を収集して改良を継続したいと考えている。音声対話を含む各種アプリケーション・フレームワーク（仕掛け）が提案されつつある。重要なのは、音声対話・誘導のための発話文生成と発話文コーパスからの音声認識環境の自動整備であると思われる。

6. 謝辞

在宅健康管理システムの音声対話ユーザ・インターフェースの研究開発にあたり、お世話になった国際医療福祉大学国際医療福祉総合研究所、日立メディコ（株）、日立情報制御システム（株）の皆さんに感謝します。

参考文献

- [1] 鳥原　： 音声対話を含むマルチ・モダル型ユーザ・インタフェースの実装とそのユーザビリティの評価の検討、情報処理学会第59回全国大会、1999
- [2] 飯田他： 在宅システムへの音声インタフェースの応用、第19回 医療情報学連合大会、1999