

視覚障害者のための Voice Windows システムによるインターネットの高度利用

2F-3

釜江常好¹、渡辺 隆行²、小出富夫³、宇野伸一郎⁴、栗原亨⁵、本多博彦⁶、田島佐和子⁷
東京大学¹、湘南工科大学²、クリエートシステム開発³、
日本福祉大学⁴、筑波技術短期大学⁵、宇宙科学研究所⁶、ITRC⁷

Use of Internet Based on a Voice-Based Platform for the Visually Disabled

T. Kamae[†], T. Watanabe, T. Koide, S. Uno, T. Kurihara, H. Honda, S. Tajima

[†]Dept. of Physics, U. of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo, 113-0033 JAPAN.

E-mail: kamae@phys.s.u-tokyo.ac.jp

1 GUIに頼らず計算機を使いこなせる Voice Windows

ユーザーとコンピューターのインターフェースがキャラクター・ベースで行なわれていたパーソナル・コンピューターの揺籃期には、多くの視覚障害者が、MS-DOSなどのスクリプト系言語でプログラムを書き、計算機がもたらす新しい可能性を実感し、知的な興奮を晴眼者と共有することができた。視覚障害が、計算機プログラム開発において、充分克服できることは、斎藤正夫が開発したスクリーン・リーダー(VDM)などに見ることができる[1]。しかし最近のパーソナル・コンピューターは、一般利用者(晴眼者)に親しみを感じてもらうため、ユーザー・インターフェースは全て、工夫を凝らしたGraphical User Interface(GUI)となってしまうている。初心者でもコンピューターを使えるようになった反面、目の不自由な人には非常に使いづらいシステムになっている。

視覚障害者が計算機を高度に利用し、プログラム開発ができる環境を整えようとする努力の、一番の成功例は、RamanのEmacspeak[2]である。彼は、GUIに対抗してAuditory User Interface(AUI)なる概念を生み出し、音声出力で複雑な情報を短時間に伝えるための方法を多く提案している。残念ながら、このシステムには、多くの制約がある。すなわち、想定されている音声合成エンジン(DECtalk)は、Windows95/98のアプリケーション・ソフトとのインターフェースが困難の上、日本語を発音することもできない。

我々は、SAPI対応の日本語音声合成エンジンが出回り始めた頃に、この状況が改善できると考え、多言語化されたEmacs(Mule[3])のWindows95/98版(Meadow[4])を音声化することに取り組み始めた[5]。また、マイクロソフトが1997年に発表しWindows98に組み込まれている、Windows Scripting Host(WSH)が、色々なスクリプト系言語で、多くのアプリケーションやシステムの操作を可能にすることも知った。WSHでは、タスク間の通信プロトコル(Active-X、COMなどと呼ばれる)を使って、Visual Basicなどで書かれたプログラムを操作し、Visual Basic for Applications(VBA)を通してWordやExcelなどのアプリケーションを操り、ファイルの操作やインターネットの利用もできる。このWSHをベースにシステムを構築すれば、晴眼者が使うGUIベースのアプリケーションを、そのまま、GUIに頼らないで操作できる可能性を示すものとして注目した[6]。ここでは、WSHを使ったインターネット利用の例と、Meadowに組み込まれた、電子メールソフトMewの例を紹介する[7]。

2 マルチタスク、マルチリンガルなテキスト音声化システム(TTS)

現在のOSはほとんど全てが、マルチタスクで処理をするので、アプリケーションは、他のタスクのことを心配しなくてもよい。しかし既存のテキスト音声化システムは、必ずしもそのように設計されていない。例えば、あるテキストを合成音声で読みながら、キー入力を音声でエコーバックさせることができない。また、我々が読み書きする技術的な文章は、ほとんど全てが、英文混じりである。優れた音声ベースのプラットフォームを作るには、まず、マルチタスク、マルチモーダル、マルチリンガルなテキスト音声化システム(合成音システム、あるいはTTSと呼ぶ)が必要となる。

我々は複数の音声(wave stream)を同時に出力する方法として、Window Driver Model(WDM)版のオーディオドライバーを使う方法と、DirectSoundを使う方法の2つを用意した。それにより、日本語TTSシステム、英語TTSシステム、漢字の詳細読みが得意なTTSシステム、キーボード読み上げ専用のTTSシステム、プログラムコードのような記号列に特化したTTSシステムなどの複数の音声システムを同時に立ち上げておき、役割分担させることもできる[8]。DirectSoundを使うTTSシステムも開発済みである。音源を三次元空間に配置することもできる。

3 WSH ベースのプラットフォーム: 基本的な部品

Voice Windows では、VBScript がもつ 3 つの基本 I/O 関数と WSH が提供する一つの基本 I/O オブジェクトを、弱視者用の GUI と AUI を付けて提供する。これらは Visual Basic 6.0 で書かれ、標準的な exe 形式とオブジェクト化された形が用意されている。WSH では、システムに関する情報などをオブジェクトとして提供されている。これらも、AUI 付で提供している。またアニメーション・キャラクターを登場させ、喋らせたり動作させることができるようになってきている。Visual Basic では、6 つの Common Dialog Box が用意されている。これらも、弱視者用の GUI と AUI を付けて、提供すべく準備中である。

4 Meadow ベースのプラットフォーム: Voice Meadow

Meadow は Windows 版の多言語対応 Emacs である。Emacs は、電子メール、Web ブラウジング、辞書検索などの多様な機能を持ち、単なるエディタを超えた優れた GNU ソフトの一つである。新しく開発した TTS サーバー [8] を用いて Meadow を音声化したものが Voice Meadow であり、これを用いれば、視覚障害者でも Meadow の優れた機能を利用できる。Voice Meadow の最初のパッケージとして、電子メールソフト Mew を音声化 (Voice Mew) している。

5 インターネットの利用

WSH では、さまざまなネットワーク設定を行ったり、ネットワークを使って他の計算機を操作することができる。この辺りは、デモの会場で実演したい。Meadow では、いろいろなネットワーク関連のアプリケーションを使うことが可能です。その中の一つである、Mew の音声化には着手している。

6 まとめと将来への展望

Voice Windows は、晴眼者が使用している高機能なアプリケーション・ソフトを、視覚障害者がそのまま利用できるようにするだけでなく、VBScript や Emacs-Lisp などのスクリプト言語を使って計算機プログラムを開発することも可能にする新しい種類のプラットフォームである。マルチタスクに対応した音声システムであるため、異なった音声システムを使う既存のスクリーン・リーダーやホームページ・リーダーを、その上で併用することも可能となる。

謝辞

本研究は、日本学術振興会の未来開拓学術研究事業「高度マルチメディア応用システム構築のための先進的ネットワークアーキテクチャの研究」(JSPS-RFTF97R16301) の援助を受けた。ここに記して謝意を表す。

参考文献

- [1] 斎藤正夫: “視覚障害者のためのソフトウェア作成”、*Journal of IPSJ*, Vol. 36, No. 12, pp.1116-1121 (1995)
- [2] T.V. Raman: “*Audio System for Technical Readings*”, PhD thesis, Cornell University (1994).
“*Auditory User Interfaces -Toward the Speaking Computer-*”, Kluwer Academic Publishers (1997).
- [3] 半田剣一、戸村哲、錦見 美貴子、高橋 直人: “Multilingual Enhancement for GNU Emacs”
- [4] 宮下 尚: “便利に使おう Mule for Windows 活用入門” 株式会社カッタシステム (1997)
- [5] 釜江常好: “Voice Meadow (音声化した Win95/98 上の Mule) 試作品” 第 4 回インターネット技術研究委員会 (ITRC) 研究会
- [6] 釜江常好: “Voice Windows の概要とデモ” 第 5 回インターネット技術研究委員会 (ITRC) 研究会
- [7] 渡辺隆行: “Meadow 上での MEW の音声化” 第 5 回インターネット技術研究委員会 (ITRC) 研究会
- [8] T. Watanabe, T. Kamae, T. Koide, H. Honda, S. Uno, T. Kurihara, S. Tajima:
“*Multilingual, Multimodal, and Multilingual Text-to-Speech System*”, submitted to *Trans. of IPSJ* (1999).