

コモン HI サービス環境の応用

2W-2

長谷川 保 杉山 博史 半田 豊 辻本 修一 知野 哲朗 永田 仁史 浦田 耕二 竹林 洋一
(株) 東芝 研究開発センター

1 はじめに

本稿ではコモン HI サービス環境 (HI ウェア) の様々なアプリケーションへの適用について述べる。HI ウェアでは、マルチメディア情報を単にビットレベルで入出力するだけでなく、情報の内容 (コンテンツ) を処理対象としてメディア変換や自然言語処理等の HI 機能を提供することを特徴とする [1]。以下では、音声認識・合成 [2]、文字認識 [3]、機械翻訳機能 [4] 等を利用した具体的な応用システム例を示し、ヒューマンインターフェースの観点から HI ウェアの効果を検討する。

2 コモン HI サービス環境の適用領域

コモン HI サービス環境は個々の情報機器システムのために開発されてきた音声認識、機械翻訳、文字認識などの HI 機能の適用範囲を、図1のように拡大し、ヒューマンインタフェースを高度化することを目的としている。

パーソナルコンピュータ、エンジニアリングワークステーション、携帯情報機器での利用や、ネットワーク対応が可能であり、多様な既存アプリケーションソフトや新規応用の場面でメディア変換処理を容易に活用できる。以下にその特徴を列挙する。

- 複数のアプリケーションから共有利用できる
- 計算機パワーの小さい計算機でも高度な HI 機能が利用できる
- API(Application Programming Interface) の統一により新しいアプリケーションの開発効率の向上 (過去の資産活用)
- 最新メディア変換処理技術をサーバ化して機能拡充できる

3 応用システム

3.1 FAX 文書の読み上げ

広く普及しているグループウェアソフト上で、複数のメディア変換処理機能を組み合わせて利用し、FAX 文書読み上げ機能を構築した (図2)。

"Applications of Common Human Interface Services Environment",
T.Hasegawa, H.Sugiyama, Y.Handa, S.Tsujimoto, T.Chino,
Y.Nagata, K.Urata and Y.Takebayashi
R&D Center, Toshiba Corp.

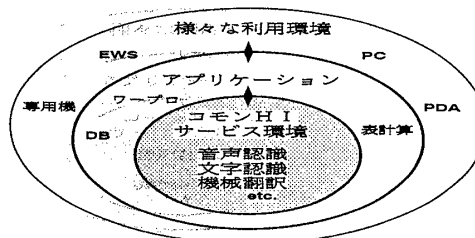


図1: コモン HI サービス環境の概念図

グループウェアに取り込まれた FAX 文書 (画像) がドキュメントリーダ (文字認識) によりテキスト情報に変換される。認識処理で文書が英文と判断された場合には、その文書は機械翻訳処理により日本語へ翻訳される。日本語のテキスト情報は音声合成部に送られ、合成音声は外部スピーカから出力される。

この FAX 文書の読み上げの指示は、図3に示すようにアプリケーションのツールバーの1つのボタンを用意し、ユーザはこのボタンを押すだけで一連の処理を実行することができる。従来 FAX の画像をドキュメントリーダで認識し、それを機械翻訳にかけ、その結果を音声合成で読ませるための複数の操作が必要だったが、HI ウェアのメディア変換連携支援により、「読みあげ」という1つのアクションで直観的な操作を行うことができる。

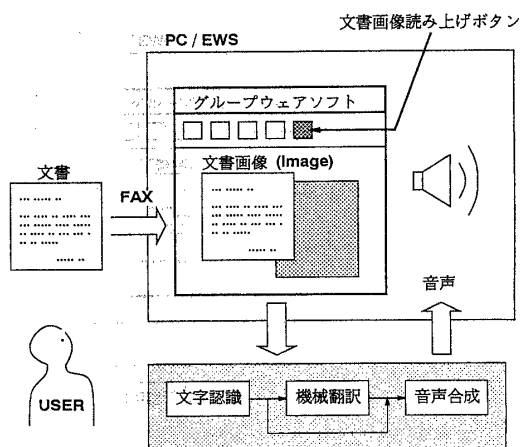


図2: FAX 文書の読み上げ

このように、HI ウェアでは各 HI 機能を既存アプリケーション上から利用できることに加え、各機能の

連携処理をボタン登録することを支援しており、複数のアプリケーションを立ち上げずに目的の処理を行うことができるという利点があり、操作が簡便で使いやすいヒューマンインタフェースの実現に役立つ。

3.2 電話音声指示電子メール読み上げ

対話的でリアルタイム性が要求される音声認識・合成機能の応用例として、電話回線を介した音声入力指示電子メール読み上げシステムを試作した。このシステムは、音声認識・合成のHI機能の他、これと直接接続できない一般のプログラムとの仲介を行う音声インタフェースマネージャと電子メールを操作するメールツール (mh-mail) とから構成される。まずモデムに電話をかけてダイヤルトーンによりログインする。システムはログインで自動的に起動し、「メールツール」という音声コマンドによりメールツールを操作対象として選択する。メールツールの操作は、着信メールのリストの中から「最後」「5番目」などの音声コマンドでメールを選択し、発信者名、サブジェクトなど自動的にエコーバックされる合成音声により目的のメールを選択したら「読む」という音声コマンドでメール内容を合成音で聞くという手順で行う。開発を通して、電話では画面表示が行えないことから、リアルタイム処理、曖昧性への対応、ヘルプ処理などが重要であることが明らかになった。

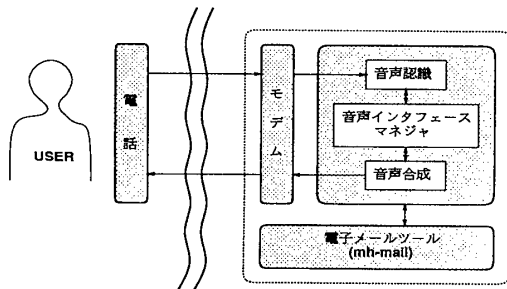


図 3: 電話音声指示による電子メールの読み上げ

本システムには外出先からでも通常の電話を介した音声入力指示により電子メールの内容を確認できるという利点がある。電子メール以外の電話帳等のデータベースシステムを電話音声入力により操作することが可能である。電話帳以外にも音声マウス入力によるマルチモーダルインタフェースも有用である。例えば CAD などの描画ソフトに用いると、音声による編集操作と、マウスによる描画操作の役割分担が可能になり、編集操作による中断なしに描画作業を行うことができる。

4 HI ウェアの使用効果

• 既存アプリケーションソフトへの対応

使い慣れているアプリケーションソフトに付加価値をつけたり使い勝手を向上させることが可能であり、メディア変換処理の品揃えを豊富にし性能を向上することにより HI ウェアの価値は高まる。

• HI 機能の連携・統合利用

各 HI 機能 (メディア変換処理機能) を互いに組み合わせ連携させることで、より複雑で高度な HI 機能が実現できる。従来「Aでの処理結果をBで処理して...」と複数のアプリケーションを操作する必要があったが、スクリプト等を用いることで1操作での処理が可能になる。このため直観的で解りやすい HI 環境を構築することができる。

• HI 機能改善へのフィードバック

様々な応用システムを開発することによりユーザ指向でメディア変換要素技術にフィードバックをかけ、操作性の向上や処理速度の向上を実現することが可能となる。

5 おわりに

本文では、音声認識・合成、文字認識、機械翻訳などの HI 機能を利用したアプリケーション開発を通して、コモン HI サービス環境の有効性を確認することができた。今後 HI 要素技術をさらに深耕するとともに、HI 機能の品揃え拡充と性能向上をはかるとともに利用環境を充実させ、新しい応用システムを構築する。

謝辞

コモン HI サービス環境の開発に際してご支援頂いた関係各位に感謝します。

参考文献

- [1] 杉山他, "コモン HI サービス環境の開発", 情処学会第 53 回全国大会, 2W-1, 1996
- [2] 橋本他, "ワークステーションにおける音声認識インタフェースの検討", 情処学会 HI 研資, HC46-3, 1993
- [3] Tsujimoto, S., "Major Components of a Complete Text Reading System", IEEE, Vol.80, No.7, 1992
- [4] 伊藤, 野村, 武田, "クライアント/サーバ型機械翻訳システムの学習方式", 情処学会第 50 回全国大会, 5R-10, 1995