

家庭用VTRにおける対話型知的インターフェースの適用

6R-4

富塚 英省、千種 康裕、福田 譲治、田村 朝子
ソニー(株) 情報通信研究所

1. はじめに

マイクロプロセッサの普及等により、様々な家電機器が高機能化している。その反面、スイッチやボタンが増えて操作が複雑になり、例えば VTRの録画予約等が行なえない人も多い。この問題を解決するため、我々はユーザと機械の間に複雑な操作を代行するエージェントを介在させた。このエージェントは、音声による命令を理解し、不足情報をユーザとの対話によって補間する。

我々は、このシステムを Super IVY(以後、SIVY)と呼び、本論文ではその概要、評価結果、及び今後の可能性について述べる。

2. 概要

従来VTRは、リモコンや本体上のスイッチ等をユーザが直接操作して、コントロールを行なう(図1)。それゆえ、命令が直接的に指定でき、素早く操作できるという利点がある反面、スイッチの数が多いため機能との対応を覚えることが困難であること、定められた手順に従って操作しなければ動作しなかったり場合によっては誤動作したりすること、等の欠点がある。

SIVYでは、画面上にキャラクタとして表現されたエージェントが VTRの操作を行う意思を持っているかの様にユーザと対話する(図2)。このようにして機械上に表現された意識を我々は、機械意識と呼ぶ。対話によるコントロールには、素早い連続操作は難しいという欠点があるが、分かり易く、操作手順を覚える必要が無いという利点がある。SIVYは自由な順序の入力を理解し、ユーザ入力の不足や誤りを対話によって補間、訂正するので様々な入力に柔軟に対応できる。また、キャラクタを用いて機械意識を視覚化し、人間同士の自然な対話状況に近付けた。これらにより、親しみ易いインターフェース環境を実現することができた。

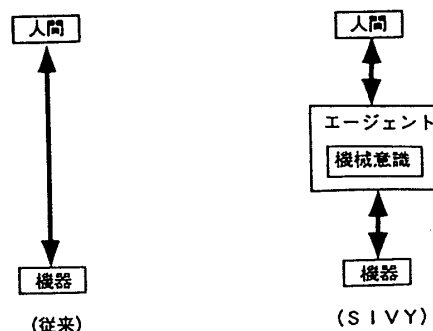


図1 インターフェース概念図

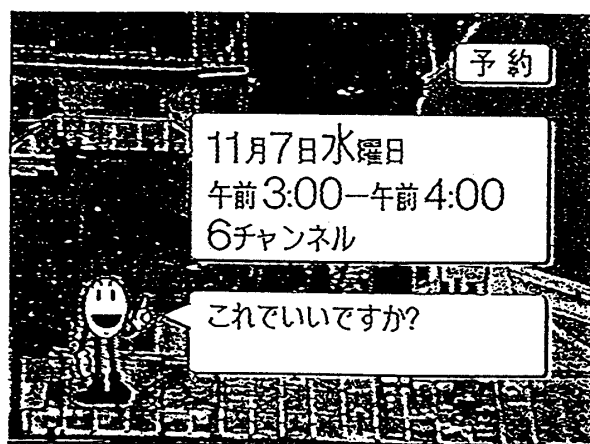


図2 画面表示例

3. システム構成

SIVYのシステム構成を、次頁図3に示す。マイクから入力された単語を音声認識部が処理し、その意味を対話理解部で解析する。その結果に応じて動作実行部がキャラクタの表情付け、動作、発声、及びVTRの制御を行なう。

音声認識部には、NAT方式音声認識装置^{[1]-[4]}を組み込み、VTRを操作する際に自然に使用されると思われる128単語を登録して使用した。機械意識の主たる機能を実現する対話理解部は、我々が以前より研究していた自然言語処理システムIVY^[5]を応用した。

Intelligent Interface for Home VTR

Hidemi TOMITSUKA, Yasuhiro CHIGUSA, Goerge FUKUDA and Asako TAMURA
Telecommunication & Information Systems Reserch Laboratories, Sony Corporation

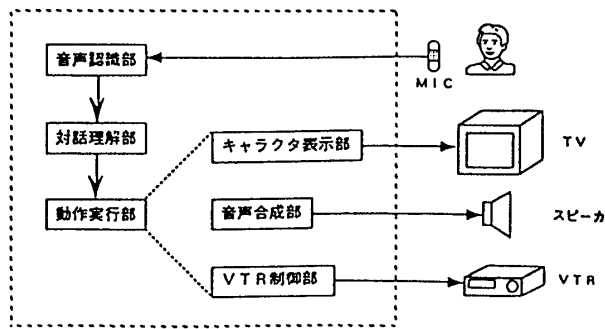


図3 システム構成図

3.1. 対話理解部

自然言語処理システムIVYは、次に挙げる三つの特徴を持つ。

- VTRの制御や予約といった限定された領域において、人間同士の対話に見られる非文法的な文をも理解できる自然言語処理。
- 常識やユーザの録画予約の習慣等を利用した不足情報の補間機能。
- 使用に慣れた人、慣れてない人といった様々なユーザのレベルに合わせられる適応的対話機能。

SIVYでは、入力時に課せられる制約を減少させるために、単語の持つ多義性や同義語の扱いも可能となるようにIVYを改良した。ユーザに親近感を持たせるため、機能的に無意味な入力、例えば感情的な表現も一部扱えるようにした。

3.2. キャラクタ動作実行部

各画面要素をグラフィックデータとして持ち、それらを逐次入れ換え表示することによってキャラクタの表情や動作を表現する。同時に、メッセージや確認内容を表示し、動作に合わせて音声合成を行なう。

メッセージの拡張等によって生じる動作パタンの追加・変更を容易にすること、及び記憶領域の節約を目的として、グラフィックデータ（コマ図）をいくつか並べた基本パターンを用意し、それを組み合わせてそれぞれのメッセージに合わせた動作を実現する手法を取った。

また、ユーザに親近感を持たせるために、ユーザの入力が一定時間行なわれなかった場合、あくびをしたり、寝たりする等、擬人的動作を行なう。

4. 評価

SIVYの操作性の評価を従来VTRと比較して行なった。

【長所】

- キャラクタとの自然言語による対話によって操作を行なうので、ユーザにとって親しみやすい。

- 1つの操作に複数の単語が用意されているため、入力の自由度が従来に比べて大きい。
- 単語の多義性も理解するので、覚える単語数が少なく済み、しかも自然に操作できる。
- 録画予約において、予約に必要な要素の入力がどのような順序で行われても正しく意味理解できた。
(例) ・日曜午後8時から45分まで1ch。
 ・NHK夜8時から45分間来週日曜。
- ユーザ入力の習慣を学習するので使用するほど入力操作简单になる。
(例) 大河ドラマを予約する習慣を学習すれば、「日曜NHK」を入力するだけで、午後8時から45分間という情報を補間できる。

【短所】

- 従来のVTRの使用に慣れているユーザは、従来方式の方が意図する操作を短時間でこなうことができる。
 - 早い反応速度を要求される操作、例えば編集等には不向き。
(例) テープを早送り状態から停止する時、発声開始からテープ停止まで約2秒要する。
- 複雑な機器を扱うことに慣れていないユーザにとって、特に有用なインターフェースであることが確認できた。

5. まとめ

機械意識を家電機器のインターフェースに適用することにより、機械に対する親しみを持たせ、複雑な操作を行なう抵抗感や手順を憶える煩わしさからユーザを開放することができた。今後は、人間科学的見地からの検討や本システムの汎用性の検証等を行なう予定である。さらに、文法知識や常識の利用等により音声認識の誤りを修復する自然言語処理の開発、複数のユーザに対する個々のモデルの設定等を行ないたい。

謝辞

本研究の機会を与えて下さった当社総合研究所 宮岡所長、情報通信研究所 松田所長に深く感謝いたします。

【参考文献】

- [1] 佐古 曜一郎 他: 新音声認識NAT, 情報処理学会第29回全国大会, pp. 1329-1330(1984).
- [2] 平岩 篤信 他: 軌跡に沿った時間軸正規化による単語認識, 日本音響学会講演論文集, pp. 17-18(1984).
- [3] 田村 震一 他: NAT方式不特定話者単語認識, 日本音響学会講演論文集, pp. 21-22(1985).
- [4] 赤羽 誠 他: 不特定話者単語認識装置の試作, 日本音響学会講演論文集, pp. 91-92(1985).
- [5] 三澤 誠一 他: 意味解析主導による対話型知的インターフェース, 情報処理学会第40回全国大会, pp. 448-449(1990).