3. 人にやさしい放送

3-3. 研究開発中の情報バリアフリー技術

正会員 浦谷 則好[†], 正会員 坂井 忠裕[†]

キーワード ● エージェント, 対話処理, インタフェース, 触覚提示, 点字端末

1. ま え が き

高齢者・障害者を含め、誰もが放送や通信の情報を享受できる「情報バリアフリー」の実現は必須の課題である。

2006年には地上デジタル放送が全国展開する。デジタル放送では、多チャンネル放送やデータ放送、電子番組表などがサービスされる。放送では幅広い利用層が対象のため、これらの情報取得にはコンピュータ以上に容易な操作で誰でも使えるアクセシビリティが要求される。しかし、現状ではリモコン操作の複雑なことや、データ放送などは基本的に視覚的な操作によるGraphical User Interface (以下、GUI) 環境のため、視覚に障害のある方が情報を得られないなどの問題点がある。ここでは、高齢の方や視覚に障害のある方がデジタル放送を享受できることを目指し、NHKが進めている情報バリアフリー技術の研究を紹介する。

2. 対話型テレビ受信システム

テレビ受信機やその周辺機器が高機能化するとともに、その操作は高齢者をはじめ、多くの人々にとっても難しいものとなってきている。誰にでも簡単に操作できるようにするためには、テレビ受信機が知的になることが必要である。すなわち、テレビ受信機がエージェント化されれば、視聴者は意図を伝えるだけで複雑な操作は受信機側ですべて代行してくれるようになる。これを実現するため、NHK放送技術研究所では、視聴者と対話できるエージェントを試作した。以下では、そのシステム構成について述べる。

2.1 テレビエージェントの構成

テレビエージェントを試作するに当たって、視聴者とシステムとの対話がどのようなかを調べるためにWOZ (Wizard of OZ: 人が機械の裏に隠れて操作し、あたかも機械が自動応答しているがごとく振る舞う方式) 実験を行った¹⁾。その結果、被験者の約83%の発話がエージェントに対する要求で、残りは自分自身への確認などの独り言であった。また、要求のうち約89%の発話には番組名、

図1 テレビエージェントのシステム構成

人名、ジャンル、時間、放送局名、返答、操作動詞等の特定のカテゴリに分類できる単語 (以下特定カテゴリー) が含まれていることがわかった。残りは、エージェントの応答の復唱や、「そうですか」などのエージェントの言葉への確認などであった²⁾。このような実験結果が得られた理由として、被験者の発話がテレビ画面上に表示される情報によって制約されたと考えられる。テレビがもっと知的なシステムになればもっと別の発話も出現してくるだろうが、被験者の心中にある現状のテレビに対するメンタルモデルも、発話の制約につながったと考えられる。

この実験結果を基に、図1のようにシステムを構成した。

(1) 音声認識モジュール・音声合成モジュール

構築したシステムでは視聴者と対話する対象を明確にするために、ぬいぐるみ型のインタフェースロボットを用いた。インタフェースロボットには超指向性マイク、スピーカ、無線装置を備えていて、音声認識モジュール、音声合成モジュールと無線で通信する。

視聴者の話した言葉は音声認識モジュールで文字列に変換される。このモジュールは、3-2章、2節で用いているものと同一で、逐次的に認識結果を確定するので対話に適したものである。ただし、テレビ操作に関する対話に適さ

[†]NHK 放送技術研究所

"Information Barrier-free Technologies under Development" by Noriyoshi Uratani and Tadahiro Sakai (NHK Science and Technical Research Laboratories, Tokyo)

図2 対話処理の流れ

せるため、言語モデルは、WOZ実験で得た発話データを用いてチューニングしている。さらに、発話に高確率で含まれる番組名、出演者名などの語彙は日々変化しているため、新しく取得した語彙をそれぞれのカテゴリに関連する言語モデルのクラスに自動登録し、言語モデルを半自動的に更新する仕組みを組んでいる。

システムからの応答は音声合成モジュールで作成される。応答音声は視聴者に与える影響の大きさについてはよく知られている。このため、インタフェースロボットの外観を考慮して、声優さんに依頼してかわいらしい声を収録してそれを用いることにした。よく使われるフレーズや有名な番組、俳優名などは録音音声をそのまま使い、新しい番組名などの登録音声がないものは、音素から合成した音声を用いるハイブリッド方式を採用している。

(2) 対話処理・プレゼンテーションモジュール

発話の内容を語の表層的な分析だけで理解するのは一般的には困難である。しかし、WOZ実験から、テレビ操作における対話に状況に限定すれば、番組名、人物名などの特定カテゴリの組合せから視聴者の発話意図を充分推定できると考えられる。そこで本システムでは、速度面での有利さも考慮して、形態素ごとのパターンマッチングで意図を推定することにした。

対話処理の流れは図2のようになる。音声認識結果に対し、形態素解析を行い、単語に品詞や属性を付与する。形態素単位でのパターンマッチングを行い、入力文のパターン、すなわち視聴者の意図を特定する。このパターンに応じて、機器制御の仕方や応答文を決定する。パターンの記述にはメタキャラクタや変数を使えるので、柔軟な記述が可能である。図2で{ }, |, *などがメタキャラクタ、@Genreなどが変数である。また、前の発話の情報も使えるように、

発話の内容はスロットで管理することによって、文脈に依存する表現にも対処できるようにしている。

エージェントの応答としては音声だけでなく、画面に情報を表示することで視聴者の利便性を高めている。また、インターネットから映画や出演者の情報を探す機能も持ち合わせている。

さらに、システムはユーザプロフィールを管理していて、「おすすめの番組を教えてください」との要求にも応えられるようにしている。そのため、テレビエージェントは、視聴者が視聴した番組や検索した番組についての履歴を視聴者毎にデータベースに保存している。視聴者が誰であるかの判断には話者認識技術や顔画像認識³⁾を利用している。

3. 目の不自由な方への情報バリアフリー端末

視覚に障害のある人の誰もが、デジタル放送でデータ放送や電子番組表 (EPG) などの情報を得られる提示装置とシステムの開発を進めている⁴⁾。

インターネット (Web) やパソコンの利用環境は、ほとんどがGUIである。晴眼者に便利なGUIは目の不自由な人には大きな情報バリアになっている。それに対し、音声読み上げや点字で情報にアクセスし、内容を得られる方式が実用化されている。また、弱視の方には、画面の特定領域の拡大や文字や背景色を反転して表示する手段や、HTML記述の文字サイズや行間隔、配色などをCSS制御で設定する方法が提案されている⁵⁾。

デジタル放送もメニュー画面から視覚的に希望の選択項目を選ぶGUIである。パソコンの利用環境と同様に、音声や点字による文字情報の伝達や、弱視の方が見やすい画面表示は不可欠である。また、データ放送のメニュー画面選択項目 (以下、選択オブジェクト) や階層の構成、EPGのよ

図3 情報バリアフリー端末と提示システムの構成

うな表形式、図やグラフなどのイメージをどのように伝えるかは重要な課題である。さらに、多くの人が接するテレビでは、パソコン以上に容易なアクセシビリティが要求される。

これらの課題の解決のために、音声や点字と併用し、空間的な認知が可能な2次元触覚提示でGUIや図などを伝える方法を検討している。

一方、視覚障害には、矯正視力が0.2以下の弱視から視野が狭い視野狭窄、全盲、視覚と聴覚の両方に障害のある盲ろうなどその幅が広い。したがって、個々のユーザがGUI環境で情報を取得するには、視覚・聴覚・触覚のすべての感覚に情報を提供する必要がある。また、視覚や触覚の特性は人により異なるので、例えば、弱視向け画面提示では反転や見やすい配色、触覚提示では弁別しやすい形状やサイズなど、個人の認知の違いに適応した多種の情報提示方式が必要である。

このような背景から、図3のように、提示に関するデータをXMLで記述し、データ放送やEPGの情報を複数の装置で多種の提示ができる方式を開発している。以下、開発中の情報バリアフリー端末と提示システムについて概説する。

3.1 文字情報を伝える音声合成、点字端末

データ放送の選択オブジェクトのタイトル情報や内容の文字情報を伝えるには音声と点字が不可欠である。音声では、人と変わらない音質の音声合成の研究を進めている。アナウンサのニュース音声素材を蓄積し、合成時には、入力された漢字かな交じり文の読みに適合する音声素材を抽出し接続する方式である⁶⁾。また、点字の初心者や盲ろう者のための点字装置を開発した。従来の点字と6指点字方式とを選択して使用できる。6指点字方式は、点字の6点を6本の指に分散して指腹に提示する⁷⁾。視覚障害者用点字ワープロや盲ろう者との会話で用いられる「指点字」と同じで、盲ろう者になじみやすい提示方法といえる。

3.2 弱視の方が見やすい画面表示⁸⁾

幅広い弱視の方の見え方に対応するためには種々の表示が望まれる。視力が低下した方には拡大表示、コントラストや配色で見やすさの異なる方には、反転表示やユーザが自由に文字や背景色を可変できる適応表示が有効である。

拡大表示では、ユーザが選択したオブジェクトの文字や図を任意のサイズに拡大できる。拡大時の文字送りは、自動的にスクロールする方法と1字、1行、あるいは1ページごとに移動する方法を備えた。

反転表示や適応表示方式では、あらかじめ設定した複数の画面から最良な画面を選択する方式と、選択オブジェクトの表示枠や文字および背景の輝度や色度を実時間で調整する方式を実現した。

拡大時の文字送りや表示画面の選択はリモコンのキー操作、文字や背景色の調整は回転型ボリュームによる容易なインタフェースとした。

3.3 GUIの画面や図を伝える2次元触覚提示端末

図3中に表すように、デジタル放送を享受できるテレビ用の触覚提示端末を目指した実験試作機を開発した。データ放送の選択オブジェクトや階層、表の理解、および情報へのアクセスのために二つの提示方式を検討し、最適な提示とアクセシビリティの評価を進めている。提示方式の一つは微小ピンアレイを平面に実装し凹凸で画面を表示する平面型触覚提示、他の一つは上下に可動な機構のスイッチを正方に実装したスイッチアレイ型触覚提示である。

平面型触覚提示では、GUIに対するタッチパネルのような機能や図やグラフなどの形状理解の支援が可能である。3mm間隔で実装できる微小ピエゾアクチュエータ(KGS製)を用い56×40の提示面を構成した。画像信号を2値化し、触覚提示の画素数に変換して凹凸で表示する。選択オブジェクトの提示には、通常の画面に忠実な提示と、配置や形状を変えて提示することができる。選択オブジェクトにア

クセスするために、ユーザが触れた指の位置をカメラで検出する方式を開発した。現在、形状認知の評価を行い、より高精度な提示方式の研究開発を進めている。

スイッチアレイ型触覚提示では、選択オブジェクトの数量や階層の深さに対応してスイッチを凸表示する。また、EPGのような表形式では行列のセルが各スイッチに対応する。ユーザは提示面に触れ、スイッチの配列や形、数量によってGUIやEPGなどの表の構成を知り、スイッチを押下して情報にアクセスできる。スイッチの配列は日常使用しパタンを認識しやすい電話のボタン配置に実装した。

3.4 XMLによる提示システムの構成

提示システムでは、各種の提示に必要なデータをデータ放送のBMLやWebのHTMLと親和性の高いXMLで記述した。XMLは拡張性のあるマークアップ記述言語である。タグの属性データを変更することで各種の装置で多種の表示が可能であり、比較的容易にワンリソース・マルチユースの構成ができる利点がある。

例えば3.2に示した弱視者の画面表示について、文字や背景色を制御するタグを<文字><背景色>とすれば、その属性データを変更することにより任意の色に設定できる。適応表示方式では実時間でタグの属性を変更している。触覚提示でも適切なタグで選択オブジェクトの配置や形状を自由に設定できる。

情報伝達のために必要なタグを抽出し、例えば「視覚障害XML」を構築すれば、情報送出と提示装置との共通なインタフェースへの可能性が広がる。

4. む す び

高齢者が簡単に操作可能な対話型テレビ受信システムと目の不自由な方への情報バリアフリー端末の研究の現状について述べた。これらのシステムは、デジタルデバイドの問題解決を図り、すべての人がデジタル技術の恩恵を

受けられることを目指している。今後も、だれもが放送を楽しんでいただけることを夢見て、研究を推進していく。

(2004年6月30日受付)

〔文 献〕

- 1) 小峯, 森田, 後藤, 浦谷: “音声インタフェースによる番組選択操作時の発話内容分析”, ヒューマンインタフェースシンポジウム2002, pp.631-634 (2002)
- 2) Jun Goto, Yeun-Bae Kim, Masaru Miyazaki, Kazuteru Komine and Noriyoshi Uratani: “A Spoken Dialogue Interface for TV Operations based on data Collected by Using WOZ Method”, ACL2003 (41st Annual Meeting of the Association for Computational Linguistics) (2003)
- 3) S. Clippingdale and T. Ito: “A Unified Approach to Video Face Detection, Tracking and Recognition”, Proc. International Conference on Image Processing ICIP'99 (1999)
- 4) 坂井, 近藤, 清水, 鈴木: “GUIに対する視覚障害者のためのマルチモーダル情報提示環境の検討”, HI学会研究会, 5, 1 (2003)
- 5) 氏間: 弱視者の使いやすいHTML文書表示ツールの試作, 第25回感覚代行シンポジウム, pp.47-52 (1999)
- 6) 世木: “高品質な音声合成”, 平成16年度技研公開予稿集, pp.46-51 (2004)
- 7) 坂井, 石原, 牧野, 近藤, 関口: “受動的触知による新しい点字伝達方式” 映情学誌, 52 (1998)
- 8) 坂井, 近藤, 鈴木, 坂尻, 岡田: “データ放送の弱視者が見やすい画面表示とインタフェースの検討”, 第29回感覚代行シンポジウム, pp.133-136 (2003)

うらやま のりよし
浦谷 則好 1975年, 東京大学大学院電気工学専攻課程(修士)卒業。同年, NHK入局。富山放送局を経て, 1979年, 放送技術研究所。以来, 情報検索, 機械翻訳, ヒューマンインタフェースの研究に従事。現在, 同研究所人間情報科学に所属。正会員。

さかい ただひろ
坂井 忠裕 1977年, 苫小牧高専電気工学科卒業。同年, NHK入局。帯広放送局を経て, 1982年, 放送技術研究所。以来, 記録技術, ヒューマンインタフェース, 視覚障害者支援などの福祉工学, 触覚の情報受容の研究に従事。現在, 同研究所人間情報科学に所属。正会員。