

住空間における音声インタフェースの利用者環境適応の検討

Personalizing speech interfaces to home environments

本間 永愛^{*1}

Hisae Honma

石川 翔吾^{*2}

Shogo Ishikawa

桐山 伸也^{*1}

Shinya Kiriyama

菟川 友宏^{*1}

Tomohiro Haraikawa

竹林 洋一^{*2}

Yoichi Takebayashi

^{*1}静岡大学大学院情報学研究科

Graduate School of Informatics, Shizuoka University

^{*2}静岡大学創造科学技術大学院

Graduate School of Science and Technology, Shizuoka University

Increased functionality in household electronics has made research on centralized user interfaces a current topic. Our research focuses on creating personalized speech interfaces to be used in home environments, where desired operations and instructions vary from one person and household to the next. We propose an environmental model where the instruction-operation correspondence is controlled according to pre-defined scenes. Each scene represents a pattern in a user's lifestyle, which we identified through data collected from real life situations. For the first stage of testing, we limited our data to test subjects who live alone with relatively stable living patterns. Based on the data, we built a test system for examining whether scene-switching according to time benefits speech interaction. Through experiments, we found our approach to increase the likelihood of a correct instruction-operation match, which results in smoother speech interaction process.

1. はじめに

近年、家電製品の高機能化やホームネットワークの技術進展とともに、生活の快適性や安全性は高まりつつある。しかし人が生活する住空間では、学校やオフィスとは違って明確な目的はなく、快適に過ごすための要求が人それぞれあると考えられる。そのため本稿では、生活する人の好みや癖、部屋や機器機能の多様性に適応するための住空間と利用者のモデリングを目的とする。筆者らの研究グループでは、ユーザーフレンドリーの面から、人にとって自然で行動の制約が少ない音声インタフェースを用い、住空間における利用者適応の環境構築に取り組んでいる。蓄積されたセンサ情報から家庭で暮らす家族の生活パターンを推測し、生活支援を行う研究 [3] や、音声インタフェースの利用者への適応研究 [1][2] がある。本稿では、利用者要求を収集・分析から、時間帯によって利用者の操作語彙や求める機器動作が異なることに着目し、時間帯情報から利用者適応を行う生活パターン適応モデルの考案を行った。また一人暮らしを対象とした適応システムの実装を行い、時間帯による切り替えを用いた利用者環境への適応について述べる。

2. 生活パターンの利用者適応

2.1 利用シーンの収集

住空間における利用者の要求や状況に対して柔軟なモデリングを行うために、利用者や環境状態のどのような利用シーンが存在するか収集・分析を行った。一人暮らしをしている人を対象に、採集者 3 人、収集期間を 1 週間とした。一人暮らしは、家電の種類をはじめ生活環境がシンプルであり、また複数人の場合を考慮する必要がないため着目した。生活に密着した利用シーンの収集を目的とし、ボイスレコーダーを用いた記録方法を採用した。生活におけるどのような利用シーンで、どの発話またはきっかけで、どのような動作（家電制御や音声対話）を行いたいかを音声で記録を行った。収集の結果、70 個の利用シーンが得られた。得られた利用シーンのサンプルを以

下に挙げる。

- 例 1. 朝, "おはよう" の発話で, 照明・テレビの電源を ON.
- 例 2. "テレビ" の発話で, 朝はニュース, 夜はバラエティにチャンネル操作.
- 例 3. 忙しい時, 確認の応答を省き音声対話の頻度を減らす.

2.2 時間帯情報による適応モデル

機器の有無など生活環境の違いによる採集者間の差違はあるものの、時間帯によって操作語彙やそれらに対応する使用したい家電機能が変化することが得られた。これらの変化は利用者の生活パターンに付随しており、主に時間帯・曜日（平日/休日）に分けられる。そのため、時間帯によって操作語彙と操作内容の対応リストの切り替えを行い、最も使いたい機能を的確に使用できる環境を実現を通し、利用者の生活パターン適応を提案する。

3. 利用者適応システム

3.1 時間を軸にしたシステム動作

時間帯による操作単語と操作内容の対応リスト切り替えの動きは、以下のタイムライン (図 1) として設計を行った。状況に合わせてパラメーターを切り替えて、操作語彙の認識辞書や語彙に対応する操作内容のリストをロードする。

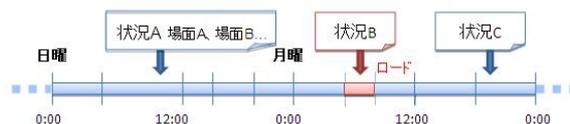


図 1: タイムラインの説明

3.2 使用例

月曜日 (平日) 朝 7 時, 環境情報をモニタリングしている状況選定部 (図 2) が, 状況 B のファイルをロードする。適応システムが利用者の目覚まし時計を止めるための「ストップ」

連絡先: 本間 永愛, 静岡大学大学院情報学研究科, 竹林研究室,
〒 432-8011 静岡県浜松市中区城北 3-5-1, 053-478-1486,
bon@takebay.net

等の発話を認識するのに備え、また対応する操作内容のリストを参照することで、家電制御を実行する。

3.3 システム構成

生活パターン適応モデルを用いた利用者適応システムの実装を行った。システム構成(図2)を以下のように設計し、対象利用者として環境や生活スタイルが固定しやすい一人暮らしの人とした。対象家電機器を、テレビ、照明スタンド、DVDレコーダーとし、操作機能を電源管理やチャンネル操作、照明の調光操作、DVDの再生操作など基本機能とした。

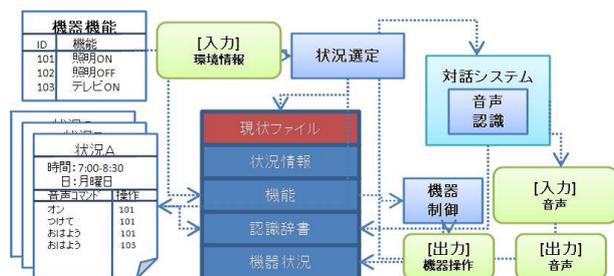


図2: システム構成図

3.4 状況設定データ構造

状況設定ファイルでは、各状況で使用される操作語彙や操作内容、状況の切り替わる条件などの記述を目的とした。状況データファイル(図3)、機器状態ファイル、操作語彙-操作機能セットファイルの3つからなる。以下に状況データファイルの記述例を示す。Eventで状況の切り換え条件を、Conditionで機器状態ファイルを参照し、Actionで認識辞書の追加・削除と単語-機器機能セットファイルを参照する。また今回これらの設定ファイルは事前に記述を行った。

```
<ScenarioControl>
  <Event>
    <Time>07:00</Time>
    <Day>Weekday</Day>
  </Event>
  <Condition>
    <State>machineState.xml</State>
  </Condition>
  <Action> //Add or Delete
    <AddGrammar>
      <Grammar>name1.grammar</Grammar>
      <Dictionary>name1.dic</Dictionary>
      <CommandSet>scenarioA.xml</CommandSet>
    </AddGrammar>
    <DeleteGrammar></DeleteGrammar>
  </Action>
</ScenarioControl>
```

図3: 状況データの記述例

3.5 音声認識

使用した音声認識エンジンは大語彙連続音声認識システム julius[4]を用いた。また、音響モデルは julius 付属の不定時話者音響モデルを使用し、音声認識では文法は記述しないものとした。利用者からの音声入力によって認識を行い、その認識結果とスコア(認識した文字の信頼度)を対話応答へ送る。これらの認識辞書は事前に登録を行った。

3.6 音声対話

閾値を定め、操作語彙とともに送られたスコアが閾値以下だった場合には、利用者にもう一度入力を促す音声応答を出力

する。逆に閾値以上だった場合には、動作確認のための音声応答を出力し、機器制御へ認識結果を送る。

3.7 機器制御

認識結果と状況設定ファイルで記述された操作語彙の比較・検索を行い、一致した操作語彙に対応する操作IDを用いて、機器機能リストから実際に操作する内容を取得する。テレビ、DVDレコーダーは、赤外線送受信機を用いてプログラムで赤外線情報の制御を行った。あらかじめ対象機器の赤外線情報を同赤外線送受信機を用いて解析し、機器機能と赤外線情報のテーブルを保持している。照明スタンドは、Dimmerを通して電圧をプログラムで制御し調光操作を行った。

4. プロトタイプシステムの評価・検討

4.1 利用者への適応性

住空間を模した室内において3章で述べた適応環境を構築し、利用者要求の変化に適応できるか検討を行った。時間帯で操作語彙や操作内容の切り替えを行うことで、生活パターンに付随した利用者の大まかな要求に適応することが出来た。また、人によって使用する操作語彙、割り当てる操作内容が異なるため、構築したシステムを用いることで、語彙と操作の多対多の適応を実現することができる。また、操作語彙と操作の対応を切り替えることで、時間の変化によって変わる利用者の要求に柔軟な対応ができる。これら操作のマクロの定義に関しては、利用者ごとに登録可能であることから適応性があるといえる。しかし現状では、事前にすべての設定を登録する必要があり大変手間であるため、これらの情報を生活の中で自動的に抽出し記述できる仕組みが必要だと考える。

5. まとめ

住空間における音声インタフェースの利用者環境への適応に向けた第一段階として、利用シーンの収集を行い、利用者状況を考慮した生活パターン適応モデルを提案した。また提案したモデルを用いて、利用者がよく使う要求に適応するシステムを設計し実装を行った。今後は、時間軸以外の利用シーンの判断要素を見つけるとともに、今回手作業で行った状況設定の記述、操作語彙の認識辞書の自動化に取り組む。また、長期使用を考えた学習アルゴリズムの開発、被験者による実験を行う。

参考文献

- [1] 守時, 荒木: 音声対話システムにおけるユーザ適応技術の統合手法の提案, 人工知能学会研究会資料, SIG-SLUD-A702-3, 2007
- [2] 吉田, 桐山, 北澤: 個人知識を考慮してカスタマイズ可能な音声対話システム, 第19回人工知能学会全国大会, 3F2-05, 2005
- [3] 美濃導彦: ユビキタスホームにおける生活支援, 人工知能学会誌, Vol.20, No.5, pp.579-586, 2005
- [4] 河原, 李: 連続音声認識ソフトウェア Julius, 人工知能学会誌, Vol.20, No.1, pp.41-49, 2005