

コモンセンス知に基づく音声インタフェース開発

Speech interface development based on commonsense knowledge

桐山 伸也*1

Shinya KIRIYAMA

*1静岡大学情報学部

Faculty of Informatics, Shizuoka University

Speech interface is not a mere 'interface' between humans and machines. Functions to understand each other by introducing internal models of the other side to both are necessary to realize smooth and smart interaction between them. In order to develop situation understanding systems based on such models, commonsense knowledge concerning each person and environment is indispensable. We have proceeded to develop speech interfaces adapting to the diversity of users' preferences, habits and rooms, and equipmental functions. The evaluation results of a prototype showed that the constructed system was useful for sustainable designing methods to adapt itself to user environment through practical use.

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、住空間の快適性や安全性は高まりつつある。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 [1] や、生活上のサービス利用を対象に、オンラインショッピングで直感的に商品検索が行える音声インタフェースの研究 [2] やポータルサイトでニュースを閲覧する際に、利用履歴からユーザにカスタマイズする研究 [3] などが盛んに行われている。家電操作をターゲットにした研究も、音声対話が可能な家電製品のヘルプシステム [4] や、指示語を用いた音声対話を実現するビデオ装置 [5]、テレビを観ながらユーザの関心があることに問い合わせが可能な対話型テレビ [6] などシステム開発が進んでいる。教師なし話者適応やハンズフリー音声認識技術により音声インタフェースの利用者適応を図る研究 [7] も進歩を遂げている。

人間と機械の円滑で気のきいたインタラクションを実現するための音声インタフェースは、人間と機械の間の単なる界面（インタフェース）ではなく、両者が相手のモデルを持ち、お互いを理解しあえる機能が求められる（図1）。そのような音声インタフェースの開発には、各ユーザの個人性を考慮することがもちろん求められるが、それと同等かあるいはそれ以上に、音声インタフェースを使用する実世界現場の状況を適切に把握すること、すなわち環境に関する知識が重要であり、ユーザ適応・環境適応は音声インタフェース研究の本質的なテーマである。このような適応型音声インタフェースの実現には、個々の状況におけるユーザの意図を理解することが求められるが、直接の入力手段であるユーザ発話音声だけから意図理解を行うことは困難である。音声入力に加えて、状況知識（＝コモンセンス知）を活用する仕組みが重要である。この観点から筆者らは、生活する人の好みや癖、部屋や機器機能の多様性に適応できる音声インタフェースの開発を進めている。

本稿では、単身者の住居環境を例題に、家電機器操作音声インタフェースを実際に使用しながらユーザの利用環境に適応させる方策の検討を題材として、個人や環境に関するコモンセンス知に基づく状況理解システムについて述べる。

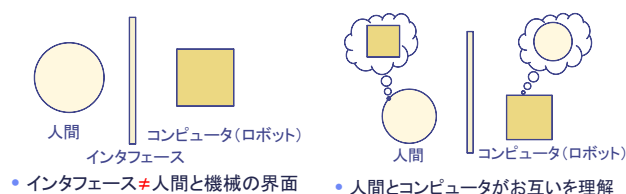


図1: 音声インタフェースのモデル

2. 住空間のユーザ環境デザイン

日常生活の拠点となる住空間では、学校やオフィスとは違って明確な目的はなく、リラックスして快適に過ごせることが重要である。世帯ごとの住人は基本的に固定であり、音声操作に使われる語彙は個人ごとに決まっておリバリエーションが少ない。操作対象の家電機器は、追加購入したり買い替えたりした際に使い方が変わるが、一度導入してしまえばしばらくは同じ機器を使い続けるのが通例である。すなわち、世帯ごと個人ごとに異なる環境間での差異は大きい、各環境内での変動は少ないというのが住空間の特徴である。

このことを踏まえ、住空間における利用者の要求や状況のモデル化を検討するため、家電機器操作場面におけるユーザ環境の具体的な事例を単身者を題材に収集した。

2.1 擬似リビング環境の構築

実環境に即した家電操作シーンの収集・取捨選択を行うため、擬似的なリビング環境を構築した。設置した家電機器は、テレビ、DVDレコーダー、照明（デスクスタンド、フロアスタンド、テーブルスタンド）、パソコンの4種類である。予備調査により、単身者の居住環境では、多くの家電操作がベッド・ソファ・椅子（作業机用）の3箇所で行われていることから、これらの家具を用意した。テレビ、照明、DVDレコーダーの3種類については、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に限定した音声インタフェースを既に開発済みであり、ユーザ環境適応のための事例収集の考察に利用した。

表 1: 単身者の住環境における家電機器操作事例集の一部

状況	時間的要素			空間的要素		操作説明	音声入力
	時間帯	曜日	季節	人	照明		
テレビを観るとき	朝	---	通年	---	---	指定したニュース番組をみる	「テレビをつけて」
テレビを観るとき	夕方	平日	通年	椅子	メイン	メインの光源を灯す	「テレビをつけて」
テレビを観るとき	夜間	休日	通年	ソファ	メイン	メインの光源の明るさを下げる	「テレビをつけて」
DVDを観るとき	---	休日	通年	ソファ	メイン	メインの光源を消す	「DVDを観る」
本を読むとき	夕方	平日	通年	ソファ	テーブルスタンド	テーブルスタンドを照射	「本を読む」
寝るとき	夜間	---	夏	ベッド	---	お休みモードでタイマーをセット	「お休みモード」
洗濯ものを乾かしたいとき	---	---	夏	---	---	除湿する	「除湿」
寝る前に音楽をきくとき	夜	平日	通年	ベッド	---	タイマーつきで音楽をかける	「音楽をかけて」
テレビを観ているとき、電話がかかってきたとき	---	---	通年	---	---	テレビの音量を下げる	「音を下げて」
カーテンを開けたとき	日中	---	夏	窓際	---	利用者に風をしっかりと当てる	「エアコンつけて」「暑い」
人が遊びにきたとき	日中	休日	夏	ソファ	---	訪問者に涼しい風が当たるようにする	「他の人に当てて」

2.2 単身者住環境における家電機器操作の事例

構築したリビング環境を活用し、家電機器操作の事例を収集した。家電を操作する状況の考察に当たっては、Open Mind Indoor Common Sense (OMICS) プロジェクト [8] の、“indoor common sense (屋内の常識知識) データベース” から、リビング環境における人間の行動や状態遷移、機器操作に関するデータを参考にした。OMICS プロジェクトでは、家とオフィス環境で利用する知的ロボット開発を狙い、屋内での人の行動や状態遷移、コンテキストに基づいた機器操作内容、部屋と物の所在関係などが世界中の人々の参加によって収集が行われている。

OMICS データベースからリビングにおける家電操作の状況を約 700 場面収集し、構築したリビング環境で実現可能な場面を取捨選択した上で、時間的要素と空間的要素に分けて状況記述項目を設計し、どの家電をどのように操作するか、どのような音声入力で駆動させるか、について整理して記述した。その結果、97 場面について家電機器操作におけるユーザ環境の事例を蓄積できた。その一部を表 1 に示す。

3. ユーザ環境の視点に基づく家電連携インターフェース

ユーザ環境適応の検討において、家電連携や機能拡張の技術が必要であるが、機器連携のフレームワークである“Molecular Home Electronics [9]”の取り組みを参考にした。ソフトウェアアーキテクチャ PIBOT (Personal Interface roBot) (図 2) では、新たに設置された家電を検出することが可能で、異なる機器の機能を連携機能として自動生成することができる。また、自動生成された連携機能の異なるビューを生成することができるため、連携機能が既設家電の機能拡張のように見え、自由度の異なる外部インタフェースデバイスとも自然に連携することが可能である。家電が追加されるたびに、個々の家電の機能を拡張するため、利用ユーザは少しずつ連携機能に慣れていくことができる。また、もうひとつの利点として、コマンドインタフェースから他のインタフェースへとモダリティ変換を行うことで、多様なモダリティへの対応を可能としている。

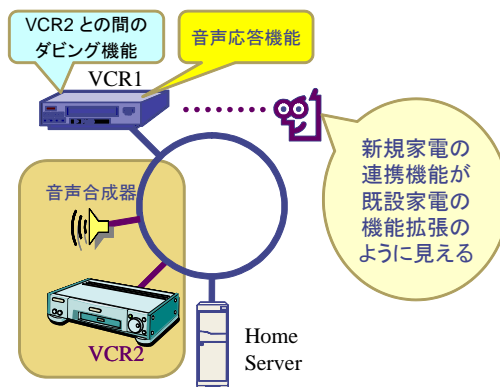


図 2: PIBOT (Personal Interface roBot) のコンセプト

4. プロトタイプシステムの構築と評価

新しいシステムを使うときに、はじめからマニュアルで全ての機能を覚える人は稀である。この観点から、必要な時に必要な機能を「使いながら」適応が進む仕組みを持つプロトタイプを設計・実装した。システムを使いながら、ユーザがシステムに熟練し、システムのユーザ状況理解が深化し、持続的にシステム機能が拡充する点が特徴である。

4.1 システムの構成

構築したプロトタイプのシステム構成を図 3 に示す。2.2 節で検討した事例から、単身者用住空間を利用する代表例である大学生に焦点を当て、家電機器を使用する生活シーンを吟味し、「テレビを観る」「音楽を聴く」「勉強をする」の 3 つを対象に、システムの詳細設計を行った。

ユーザが使いながらシステムをパーソナライズできる点が本システムの特徴である。テレビ画面を情報提示装置として利用し、各々の時点でユーザの利用可能性が高い機器操作リストを提示する。ユーザは提示されたリストを参照しながら機器操作に対応するコマンドを発話し、単独ないし複数の家電機器を操作できる。ユーザの操作履歴を保存し、利用した時間帯の情報と、家電機器の状態を手掛かりとして、動的に機器操作コマンドリストを生成する。

初回操作時のシステム状態遷移 (図 4) と、使用履歴に基づいて適応した後のシステム状態遷移 (図 5) の一例を示す。

音声認識部

音声認識エンジンには大語彙連続音声認識システム julius [10] を用いた。また、音響モデルは julius 付属の不定話者音響モデルを使用し、単語音声認識用の文法を記述した。操作語彙の音響尤度が閾値以下の場合、ユーザに再入力を促す音声応答を出力し、閾値を越える場合には、動作確認のための音声応答を出力するとともに、機器制御部へ認識結果を送る。

機器制御部

音声認識結果と現状記述ファイルにより操作語彙の比較・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と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。

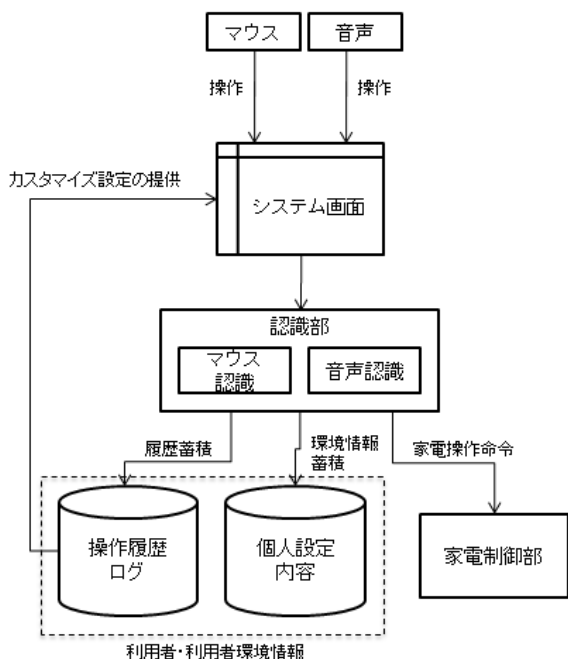


図 3: システム構成

4.2 システムの評価

構築したりピン環境で前節で述べたシステムを稼動し、状況に即した機器操作コマンドをナビゲートする機能の有効性について大学生を対象に検討を行った。日常的によく経験するシチュエーションに即した案内をしてくれる点は概ね好評であり、スイッチのオンオフのように物理的な機器操作のレベルではなく、目的指向でナビゲートする方策が受け入れられた。使用履歴に基づいて次回以降の操作がショートカット的に容易になる機能についても効果がみられた。まとめて複数の機器を操作する場面がもっと増えるとよい、音声入力の語彙をもっと柔軟にカスタマイズしたいなどシステム改良に向けたコメントを多数獲得できた。操作履歴を参照したところ、デフォルトで設定した以外の複数機器の連携操作などカスタマイズのための新たな活用方法のヒントも得られた。

5. おわりに

個人や環境に関するコモンセンス知を搭載した音声インタフェースの検討について述べた。住空間における家電機器操作音声インタフェースを、ユーザの利用環境に適應させる方策を検討した。単身者の住居環境における一般的な家電機器群を対象に、時間帯、家電機器の状態、ユーザの使用履歴を考慮して状況に即した機器操作をナビゲートする仕組みを実現した。大学生の生活環境での使用を想定したプロトタイプシステムを構築・評価し、実環境で使いながらユーザ環境適應デザインを持続的に進められる見通しを得た。

参考文献

[1] 美濃 導彦: コ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.

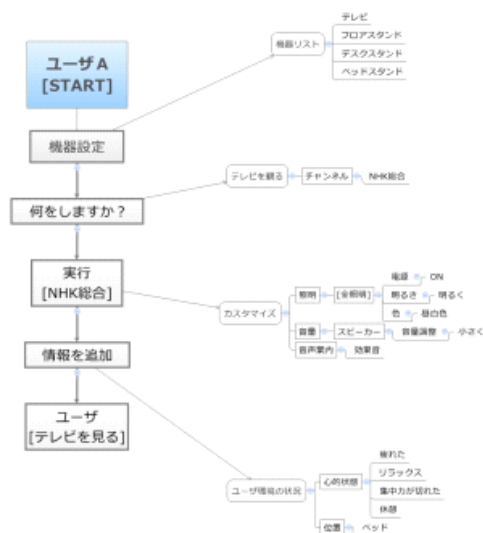


図 4: ユーザ環境適應前のシステム状態遷移

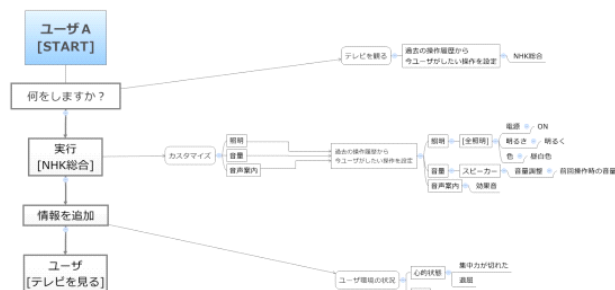


図 5: ユーザ環境適應後のシステム状態遷移

[2] 西本 卓也, 岩田 英三郎, 櫻井 実ほか: 探索的検索のための音声入力インタフェースの検討, 情報処理学会研究報告, Vol. 2008, No. 11, pp.9-14 (2008).

[3] 藤井 崇史, 平田 敏之, 三浦 元喜ほか: インターフェースをパーソナライズするニュース提供システムの提案, 情報処理学会研究報告, Vol. 2006, No. 34, pp.13-18 (2006).

[4] 伊藤 亮介, 駒谷 和範, 河原 達也: 家電製品のマニュアルの知識と構造を利用した音声対話ヘルプシステム, 情報処理学会研究報告, Vol. 2001, No. 68, pp.1-6 (2001).

[5] 滝口 哲也, 有木 康雄, 佐古 淳: コンテキストウェアネスに基づく対話型テレビの検討, 情報処理学会研究報告, Vol. 2005, No. 103, pp.25-30 (2005).

[6] 渡辺 裕太, 関口 芳廣, 鈴木 良弥ほか: ビデオ装置を例とした家電品の音声対話機能について, 情報処理学会論文誌, Vol. 44, No. 11, pp.2690-2698 (2003).

[7] 鹿野 清宏: 人に優しく環境にも適應する音声インタフェースの研究開発, システム/制御/情報, Vol.49, No.2, pp.53-57, 2005.

[8] Open Mind Indoor Common Sense: <http://openmind.hri-us.com/login.jsp>

- [9] T.Haraikawa, S.Nakagawa, Y.Une, Y.Sakane, Y.Takebayashi: A unique way to personalized and situational supports for universal access of home appliance network, Consumer Electronics, 2006. ICCE apos;06. 2006 Digest of Technical Papers. Internation Volume, issue 7-11 pp.59-60 (2006).
- [10] 河原 達也, 李 晃伸: 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.