

暗黙的なユーザ要求を抽出・推定する ホームネットワークのための対話型音声インターフェース

松原 典行[†] 江上 公一[†] 井垣 宏[†] 中村 匡秀[†]

[†] 神戸大学 〒657-8501 神戸市灘区六甲台町 1-1

E-mail: †{matsubara, egami, igaki, masa-n}@ws.cs.kobe-u.ac.jp

あらまし 多様なネットワーク家電やセンサ機器が数多く接続されているホームネットワークシステム (HNS) では従来の家電で用いられてきた専用リモコンに代わる, 新しい操作法が求められている. 音声による機器制御を実現する音声インターフェースはその中の一つである. 特に混合主導型音声対話システムは, ユーザの学習コスト削減などのメリットから, 今後の普及が期待されている. 一方で, HNS にこの種の音声インターフェースを適用するには, (1) ユーザの暗黙的な要求を操作へ反映できない, (2) 複雑な対話パターンの実装が必要なため, 開発コストが増大する, (3) 音声インターフェースからのフィードバックの無駄が多い, といった問題点がある. 本稿では, ユーザの発話内容や家電・環境の状態からユーザの暗黙的な要求を抽出・推定し, 家電機器を操作する対話型音声インターフェースを提案する. 開発した音声インターフェースによる実験では, 暗黙的なユーザ要求に対応し, 適切に機器操作を行うことが可能であることが確認できた.

キーワード ホームネットワークシステム, 音声認識, 音声リモコン, 対話型インターフェース, 暗黙的な要求

Interactive Voice Interface for Eliciting and Estimating Implicit User Requirements in Home Network System

Noriyuki MATSUBARA[†], Kouichi EGAMI[†], Hiroshi IGAKI[†], and Masahide NAKAMURA[†]

[†] Kobe University rokkoudaityou 1-1, nada-ku, Kobe, Hyogo, 657-8501 Japan

E-mail: †{matsubara, egami, igaki, masa-n}@ws.cs.kobe-u.ac.jp

Abstract In Home network system, various appliances, sensors and user interfaces are connected to the home network. Novel user interfaces are expected to control such appliances, instead of conventional remote controllers. Voice user interface is one of novel user interfaces. Especially, a mixed-initiative voice interaction system is expected to help effective appliance control with users. However, there are several problems in application of such mixed-initiative voice interaction in HNS. First, implicit requirements to appliances by users are not considered. Secondly, complex and sophisticated voice interaction patterns must be implemented to realize such voice interaction system. Furthermore, in many cases, the mixed-initiative voice interaction system bore users with redundant feedback. In this paper, we propose interactive voice interface in HNS to elicit and estimate implicit user requirements. We indicated usability of our voice interface to control appliances with implicit requirements by users through some comparative experiments.

Key words Home Network System, voice recognition, voice remote controller, interactive voice interface, implicit requirements

1. はじめに

ホームネットワークシステム (HNS) は, 宅内の様々な家電やセンサをネットワークに接続し便利で快適なサービスを実現する次世代のユビキタスアプリケーションの枠組みとして盛ん

に研究されている [1]. 家電をネットワークに接続することで, 宅内外からの遠隔監視や遠隔操作, さらには複数機器の連携や外部ネットワークリソースとの融合など, 様々な付加価値サービスの登場が期待されている [2].

HNS における重要な研究課題のひとつとして, ユーザイン

ターフェース (UI) が挙げられる。HNS はその構成機器やユーザ層の多様性から、従来家電で用いられてきた個別機器の専用リモコンや統合操作パネルで制御するには限界があると考えられている [3]。音声インターフェースは最近注目されつつある新しい UI の一つである。音声によって家電機器を制御するという手法は既にいくつか提案されており [4]、一般的にユーザが発話する特定のコマンド (音声コマンド) を機器制御のための API と対応づけることで、機器の操作が実現される。音声コマンドと API の対応づけによって音声インターフェースの振る舞いが決定されるため、異なる HNS 環境への導入が比較的容易である。また、発話するだけでよいので、多様なユーザが容易に利用できる。

一方で、ユーザは機器を利用するための全ての音声コマンドを学習 (記憶) し、適切に発話しなければならない。そのため、利用する機器・サービスが増大するにつれて学習コストは増大する。また、ユーザの家電制御コマンドそのものに対する要求 (明示的な要求) は満たされるが、具体的な家電制御コマンドを含まないような要求 (暗黙的な要求) は満たされない。

そこで本稿では、暗黙的なユーザ要求抽出・推定、学習コストの削減とそれに伴う弊害の低減、を目的とした HNS のための対話型音声インターフェースを提案する。

対話による暗黙的なユーザ要求抽出・推定： HNS において、家電制御に関するユーザ要求は明示的なものと暗黙的なものが存在する。明示的なものとは、家電の制御コマンドそのものに対する要求である。また暗黙的なものとは、直接的には家電の制御コマンドを指定せずに行われる家電制御についての要求を指す。一般的に、暗黙的な要求はユーザがおかれている現在の状態に対する不満や家電の制御後の理想的な状況に対する期待といった形で提示されることが多い。例えば、「暑い」という発言は現状に対する不満であり、「冷房もしくは扇風機をつけたい」といった要求を暗黙的に含んでいると抽出・推定することが可能である。そこで本研究では、音声インターフェースとの対話によって出てきた暗黙的な要求を抽出・推定し、家電制御コマンドへと変換する仕組みを実現した。

テンプレートメソッドパターンを用いた混合主導型対話の実現：ユーザが学習すべき音声コマンドの量を減らす試みとして、混合主導型対話 [5] が提案されている。この手法ではユーザが学習していない音声コマンドについては、音声インターフェースがフィードバックする選択肢から選択するなど、学習コストの削減が可能となっている。一方で、事前に個々の対話シーケンスを全て定義しておく必要があり、開発時のコストが非常に大きくなる可能性がある。そこで我々は HNS において有効な対話パターンをテンプレート化し、混合主導型対話システムを容易に実現するための枠組みを構築した。

家電状態を利用したコマンド提示：家電機器は通常、状態を持ち、予め定められた状態遷移モデルに基づいて、その状態を遷移させる (ex. 電源状態の ON OFF)。HNS では一般的に、ユーザが家電を制御する際には、現状態とは異なる状態へと遷移させることを目的とする。そこで我々は、家電の現状態と状態遷移モデルを音声インターフェースに取り込むことで、現在

実行可能な操作の絞り込みを行い、その絞り込みによって、音声インターフェースが提示するフィードバック内容を効率化し、長大化を防止する仕組みを導入した。

また、本稿では、音声認識エンジン Julius4.1.2 [6] を用いて、提案法の実装を行い、現在運用中の HNS (CS27-HNS) [2] に組み込み、評価実験を行った。11 名の被験者による評価実験では、全被験者が適切に音声を用いて機器を操作することができた。また、そのうち 6 名の被験者については、暗黙的な要求の発話によって被験者の期待した状態に機器を制御することが可能であることが確認できた。

2. 準備

2.1 ホームネットワークシステム (HNS)

HNS は宅内の扇風機・エアコン・照明・カーテンなどの家電機器や、温度計・照度計などのセンサと、それらを制御するホームサーバから構成される。HNS 内における家電機器はネットワーク家電と呼ばれ、ユーザや外部エージェントがネットワーク越しに制御できるように、Application Program Interface (API) を公開しており、この API を通じた機器制御や機器動作状態の獲得が可能である。実際に我々が現在運用している CS27-HNS [2] では、扇風機・エアコン・照明・カーテンといった家電機器がその機器機能を Web サービス API として公開しており、多様なインターフェースを通じて制御できるようになっている。

2.2 HNS における機器操作の課題

HNS 機器の操作インターフェースとして代表的なものに各機器の専用リモコンがある。個別機器の専用リモコンは HNS 環境だけでなく、従来の様々な家電に対して最もよく用いられてきたインターフェースである。これらのインターフェースは機器に備わった機能の全てを利用できるが、操作対象によってはボタンの数が増え、操作が煩雑になってしまうものもある。

このようなリモコンを用いたインターフェースの場合、ユーザは操作したい機器のリモコンの仕様を学習しなければならない。そのため、HNS に配備される機器の種類や数が増加すると、ユーザの学習コストが増加し、操作が煩雑になる傾向がある。特に、多くの高齢者や障害者はリモコンを操作する際に、リモコンの液晶表示やボタンラベルの視認がしづらい、機能やボタンの数が多すぎて操作が難しい、といった不満を感じている [3]。

一方で、健常者が利用する際でも、状況によっては不便だと感じることもある。例えば、室内が暗くてリモコンの位置が分かりにくい、料理中で手がふさがっていて操作ができない、靴を履いてから家電機器の切り忘れに気づき、また靴を脱いでリモコンのところまで行かなければならない、などといった状況が挙げられる。

また、従来の専用リモコンでは、ユーザの要求は家電機器を直接操作することでしか機器操作に反映させることが出来ない。よって、例えば、ユーザが「暑い」といった暗黙的な要求を持ったとしても、それをそのまま機器操作に反映させることはできず、結局ユーザ自らが空調機器を操作しなければならない。

以上のような点から、リモコンやパネルといった特別なデバ

イスを利用する必要がないインターフェースが現在期待されている。

2.3 従来の音声による HNS 機器操作

操作のための特別なデバイスを必要としない次世代の UI の一つとして、音声インターフェースが注目されている。音声インターフェースでは、音声認識エンジンを利用してユーザの発話内容を認識し、認識結果と特定の機器制御のための API を対応づけることで、機器の操作が実現される。このため、異なる HNS 環境への導入が比較的容易なインターフェースである。

家電を対象とした音声インターフェースとしては、ビデオを録画する操作を音声で行うもの [4] や頭の向きと音声を併用した四肢まひ者向けの家電機器操作インターフェース [7] などが提案されている。

2.4 音声インターフェースの HNS への適用可能性

既存の HNS 向けの音声インターフェースの多くでは、発話された内容を音声コマンドとして解釈し、機器制御のための API と対応づけることで、機器操作を実行する。

このとき、音声インターフェースの実現手法として、一度の発話の中に操作する機器名、実行する操作名、操作の実行に必要なパラメータ名の 3 つの情報を要求する手法と、インターフェースとの対話により段階的に機器操作に必要な情報を取得する混合主導型対話 [5] を用いた手法がある。混合主導型対話の最大のメリットはユーザが必ずしも全ての音声コマンドを事前に記憶している必要が無いという点である。本稿においても、ユーザの音声コマンド学習コストを削減する目的で後者の手法を採用している。この混合主導型対話を用いた音声インターフェースを実現する上で考慮すべき問題点は、以下の三つが考えられる。

P1: ユーザの暗黙的な要求の抽出・推定の困難さ 一般にユーザが家電操作に際して行う要求には、操作対象の機器が決まっている明示的な要求と、明確には操作対象機器が定まっていない暗黙的な要求の 2 種類が存在する。明示的な要求では、ユーザはどの機器をどのように操作すべきかを明確に理解しており、そのための機器操作コマンドを指示することができる。暗黙的な要求では、ユーザは機器操作そのものではなく、機器操作を実行した結果生じる状態に関するイメージしか持っておらず、必ずしも具体的な機器操作コマンドを指示できない。

P2: 対話パターンの増大による開発コストの増加 混合主導型対話システムでは、想定されるユーザとの対話パターンごとに、処理を切り分ける必要がある。そのため、家電機器数の増加・機器操作数の増加などで想定されるユーザとの対話パターンが増大するに伴い、音声インターフェースの開発コストが増加してしまう。

P3: 音声インターフェースによるフィードバックの長大化 混合主導型対話ではユーザが操作すべき機器操作コマンドに関する情報を音声インターフェースがフィードバックする。しかしながら、実行可能な機器操作やパラメータが多ければ、それだけ音声によるフィードバックが長大化してしまうという問題点が存在する。

次章では、これらの問題点の改善を目的とした本研究で、我々

が提案する音声インターフェースについて詳述する。

3. 提案法

3.1 キーアイデア

本稿では、前節で述べた問題点の改善を目的として、以下に示すキーアイデアを提案する。

K1: 対話による暗黙的なユーザ要求抽出・推定 本稿では、暗黙的なユーザ要求を温度や照度といった環境に対するユーザの不満や期待として抽出する。ここで抽出された不満や期待を環境種別ごとに分類することで、対応する機器操作の推定を行う。例えば、「暑い」という温度に関する不満を暗黙的なユーザ要求として抽出し、温度に影響を与える家電機器から適切な機器操作の推定を行う。

K2: テンプレートメソッドパターンを用いた混合主導型対話機構の実現 テンプレートメソッドパターンは、特定の処理のおおまかなアルゴリズムを予め決めておき、詳細の具体的な設計を必要に応じてサブクラスに任せるとを目的としたオブジェクト指向ソフトウェアにおけるデザインパターンである [8]。我々は音声操作のための対話パターンをテンプレートとして設計することで、多様な機器、多様な対話パターンへの音声インターフェースの柔軟な適応を実現した。

K3: 家電状態を利用したコマンド提示 HNS に配備される家電はそれぞれ内部状態を持っており、その各状態で実行可能な操作や必要なパラメータが決まっている。我々は音声インターフェースにおける家電の内部状態に着目することで、現在実行可能な操作の絞り込みを行う。絞り込みにより、音声インターフェースが提示するフィードバック内容の削減を実現する。

3.2, 3.3, 3.4 節ではこれらのキーアイデアの詳細について述べる。

3.2 対話による暗黙的なユーザ要求の抽出・推定

本稿では、暗黙的なユーザ要求を環境に対する不満もしくは理想的な状況に対する期待として抽出する。ここで環境とは、家電機器の実行によって変化するユーザを取り巻く状況を指し、温度や照度、音といった種類の要素が存在する。

本稿では、温度、照度、音の各要素を HNS においてユーザが抱くであろう暗黙的な要求の対象とする。これらの環境要素に対する不満や期待といったユーザの発話内容を環境要素ごとに分類し、音声認識を行うことで、ユーザの暗黙的な要求がどの要素に対する要求であるのかを抽出する。

以下は環境要素ごとの暗黙的な要求に関するユーザの発話事例である。

- 温度: 「寒い」「暑い」およびその類義語
- 照度: 「暗い」「明るい」「まぶしい」およびその類義語
- 音: 「音楽」「静かだ」「うるさい」およびその類義語

「暑い」というユーザの発話を認識することで、温度に対する暗黙的な要求であると判断することができる。環境要素ごとの暗黙的な要求の抽出を行ったあと、その要求が具体的にどの機器に対する操作要求と対応づけられるかを推定する。そのため我々はまず、HNS に存在する機器操作を、どの環境要素にどのような影響を与えるかという観点で以下のように分類し

た．ここでは操作対象機器を，扇風機，エアコン，ライト，電動カーテン，オーディオ，の5種類とする．

- 温度を上げる：扇風機を消す，エアコン（暖房）をつける，エアコン（冷房）を消す
- 温度を下げる：扇風機をつける，エアコン（冷房）をつける，エアコン（暖房）を消す
- 照度を上げる：ライトをつける（日中の場合）電動カーテンを開ける
- 照度を下げる：ライトを消す（日中の場合）電動カーテンを閉める
- 音を出す：オーディオをつける，オーディオのボリュームを上げる
- 音をなくす：オーディオを消す，オーディオのボリュームを下げる

これらより，HNS においての暗黙的要求に関するユーザの発話を具体的な機器操作とつなげることが可能となる．その例としては，以下の様なものが挙げられる．

- 「寒い」およびその類義語：上記の『温度を上げる』の項目にある各操作
- 「暑い」およびその類義語：上記の『温度を下げる』の項目にある各操作
- 「暗い」およびその類義語：上記の『照度を上げる』の項目にある各操作
- 「明るい」「まぶしい」およびその類義語：上記の『照度を下げる』の項目にある各操作
- 「音楽」「静かだ」およびその類義語：上記の『音を出す』の項目にある各操作
- 「うるさい」およびその類義語：上記の『音をなくす』の項目にある各操作

この様に HNS において発言されるであろう現状に対する不満や言及を暗黙的な要求とすることで，機器操作コマンドへの変換が可能になり，キーアイデア K1 を達成することができる．

3.3 テンプレートメソッドパターンを用いた混合主導型対話機構

本稿で提案する音声インターフェースは混合主導型対話システムとして実現される．そのため，音声インターフェースは対話のパターンに応じた複雑な対話シーケンスを実装する必要がある．そこで我々は，テンプレートメソッドパターンに基づき，図1に示す音声インターフェースの標準対話テンプレートを作成した．標準対話テンプレートには，*input*，*recommend*，*select*，*invoke* という対話処理コンポーネントが存在する．

この標準対話テンプレートを利用し，明示的な要求と暗黙的な要求の2種類の対話パターンを作成した．以降では対話パターンごとの対話処理コンポーネントの振る舞いについて説明する．

明示的な機器操作要求パターン

明示的な機器操作要求パターンでは，ユーザが操作したい機器操作コマンドを発話する．通常，ユーザは「機器名」，「操作名」，「操作時のパラメータ」の全てを発話する必要がある．しかし，混合主導型対話機構を用いることにより，ユーザは音声

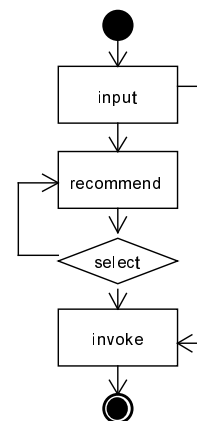


図1 標準対話テンプレート

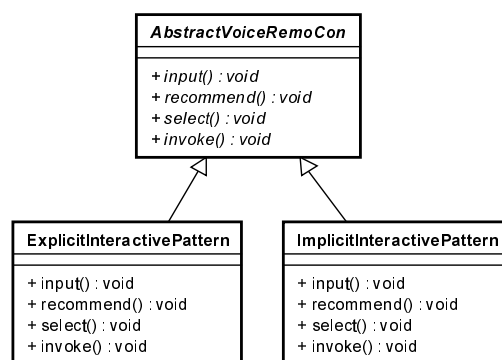


図2 対話パターンクラス

インターフェースがフィードバックする機器操作コマンドを順に指摘するだけで機器操作が可能となる．このパターンのシーケンスは以下に示すとおりである．

input：ユーザが機器操作コマンドを発話する．ユーザが発話する機器操作コマンドは(1)「機器名」のみ，(2)「操作名」のみ，(3)「機器名」と「操作名」，(4)機器操作コマンド全て，の4パターンが存在する．この中で(1)，(2)，(3)が発話された場合は *recommend* に処理が遷移し，(4)の場合には *invoke* に遷移する．

recommend：音声インターフェースによる機器操作コマンドの推薦が行われる．推薦内容として *input* でユーザ発話の内容に応じて，機器操作コマンドの完成に必要な情報がフィードバックされる．音声インターフェースによるフィードバック後，*select* に遷移する．

select：*recommend* で音声インターフェースがフィードバックした内容をユーザが選択し，発話する．ユーザの発話によって機器操作コマンドが完成した際には *invoke* に遷移し，未完成の場合には *recommend* に戻る．

invoke：完成した機器操作コマンドを実行する．

暗黙的な機器操作要求パターン

暗黙的な機器操作要求パターンは3.2節で述べたユーザからの暗黙的な要求の抽出と機器操作の推定を対話により実現する．

input：ユーザは暗黙的な要求として環境に対する不満や期待を発話する．ここでは「暑い」「寒い」といった暗黙的な

要求が発話される．発話後に *recommend* が実行される．

recommend : *input* で発話された暗黙的な要求がどの環境要素をどのように変化させることを望んでいるのかを推定する．推定結果に基づき，暗黙的な要求に合致する機器操作のリストをユーザにフィードバックし，*select* に遷移する．

select : ユーザが *recommend* でフィードバックされた機器操作のリストからユーザの望む機器操作を選択し，発話する．その後 *invoke* に遷移する．

invoke : *select* で指定された機器操作コマンドを実行する．標準対話テンプレートに基づいて作成した上記 2 つの対話パターンクラスを図 2 に示す．本稿で述べた標準対話テンプレートを利用することで，複数の対話パターンの音声インターフェースへの導入が容易に実現可能になった．

3.4 家電の状態に応じたコマンド提示

HNS に配備される家電にはそれぞれ内部状態があり，状態ごとに実行可能な操作，および必要なパラメータが決まっている．例えば，エアコンは電源の状態を持っている．電源がオフの状態では，温度を上げたり下げたりする操作は実行できず，電源をオンにする操作だけが実行可能である．また，一般的な自動カーテンであれば，「開」「閉」という二つの状態を持っている．「カーテンを開ける」という操作は，カーテンの状態が「閉」の時のみ実行可能で，「カーテンを閉める」という操作は状態が「開」の時のみ有効である．このように各家電機器が現在状態で実行可能な機器操作は，その機器が保持する全ての機器操作より少なくなる．

我々はこれを利用し，3.3 節で述べた *recommend* において音声インターフェースがフィードバックする内容を短縮した．すなわち，ユーザが現在実行できる機器操作のみを音声インターフェースがフィードバックすることで，フィードバック内容が無駄に長大になることを防いだ．

3.5 実装

提案した 3 つのキーアイデアに基づいて，HNS 機器を操作するための音声インターフェースシステムを実装した．提案システムの実装環境を以下に示す．

音声認識エンジン：音声認識エンジンには Julius4.1.2 [6] を用いた．Julius は認識すべき語彙や文法を自由に定義できることを特徴としている．本稿では，文法 7 種類，語彙 52 種類を登録した．

音声合成エンジン：フィードバックを出力するための音声合成エンジンには，PENTAX 社製の VoiceText Engine を用いた．

実装言語：音声認識エンジンを利用するためのアプリケーションを Eclipse 上で Java 言語で実装した．

3.6 操作・実行例

ここで 2 つの操作・実行例を挙げる．

・明示的な要求の場合（初期状態：扇風機 = オフ）

ユーザが「扇風機」と言うと，家電操作のテンプレートが呼び出され，まず *input* が実行される．次に *recommend* に移り，システムは「扇風機の何を実行しますか？利用できる操作はオンです．どうぞ」と推薦する．次に *select* に移り，ユーザが「オン」と選択する．すると *invoke* に移り，扇風機のオンが実

行される．

・暗黙的な要求の場合（初期状態：冷房 = オン，扇風機 = オフ）

ユーザが「暑い」もしくはその類義語を言うと，温度に関するテンプレートが呼び出され，まず *input* が実行される．すると *recommend* に移り，システムは「それでは扇風機のオンを実行しますか？それとも何もませんか？どうぞ」と推薦する．次に *select* に移り，ユーザが「扇風機」もしくは「オン」を含む発言で選択する．すると *invoke* に移り，扇風機のオンが実行される．

4. 評価実験

4.1 実験の概要

提案手法の有効性を評価するために，被験者に提案システムを用いて HNS に接続された家電機器を操作してもらう実験を行った．本実験では，部屋の環境をこちらで指定した状態（以下，理想状態とする）にするタスクを被験者に行ってもらった．ここで理想状態とは，照明が全灯である，カーテンが開いている，暖房がついている，オーディオが再生されているという状態を示す．また，このタスクを行ってもらう際は，全ての家電の電源がオフ（電動カーテンはクローズ）状態になっているものとする．本実験では音声インターフェースの使用経験がない，もしくは日常的に使用していない 11 名（20～30 代の男性 10 名，女性 1 名）を被験者として実験を行った．

4.2 実験環境

対象とする HNS として，現在運用中の CS27-HNS [2] を用いた．操作対象とした機器は，扇風機，エアコン，ライト，電動カーテン，オーディオの 5 種類である．音声の入力装置としては，Plantronics 社製 Bluetooth ヘッドセット Voyager510 を用いた．

4.3 実験手順

以下に実験手順を示す．

1. 被験者に実験内容の背景（HNS など）の説明と実験に関するインフォームド・コンセントを得る．
2. 提案システムの基本的な操作方法の説明を行う．
3. 実験の流れとタスクが書かれた用紙を被験者に渡し，それについて説明を行う．
4. 提案システムを用いてタスクを行ってもらう．

なお，2. で被験者に説明した内容は以下の通りである．

- ・ 機器名を発話すれば，基本的な操作はできるようになっている．利用可能な機器を知りたいときは「家電ピックアップ」と発話すればよい．
- ・ 環境（温度，照度，音）に関して感じたことを発話してもよい．
- ・ 「音声コンピュータ」と発話すれば，提案システムの起動と終了を行える．
- ・ 「キャンセル」「何もしない」などの否定的な内容を発話すれば，システムが起動直後の状態に戻る．
- ・ 提案システムが最後に「どうぞ」と言うまで，被験者は発話しないようにしてほしい．

表 1 (A) 暗黙的な要求を含む発話を行った被験者の実験結果表

被験者No.	1	2	3	4	5	6
年代	20代	20代	20代	20代	30代	20代
性別	男	男	男	男	男	男
M1:総発話回数	24	13	39	15	30	16
M2:総発話認識成功回数	15	11	17	11	15	9
M3:暗黙的な発話認識成功回数	1	2	6	3	3	2
M4:総発話認識失敗率(%)	37	15	56	26	50	43
M5:操作完了時間(秒)	336	194	429	232	316	219
M6:フィードバック削減文字数	0	46	44	46	17	23

表 2 (B) 明示的な要求のみを発話した被験者の実験結果表

被験者No.	1	2	3	4	5
年代	20代	20代	20代	20代	20代
性別	女	男	男	男	男
M1:総発話回数	16	27	12	13	19
M2:総発話認識成功回数	15	12	12	12	12
M4:総発話認識失敗率(%)	6	55	0	7	36
M5:操作完了時間(秒)	213	270	200	216	252

4.4 実験結果

実験結果を表 1 と表 2 に示す。表 1 は発話に暗黙的な要求が含まれていた被験者の結果を、表 2 は発話が明示的な要求だけであった被験者の結果を示している。またこれ以降、発話に暗黙的な要求が含まれていた被験者群を (A)、発話が明示的な要求だけであった被験者群を (B) とする。

指標として以下のメトリクスを収集した。計測対象としたのは、被験者がシステムを起動してから初めてシステムが「どうぞ」と言ってから、被験者がシステムを終了するために「音声コンピュータ」と発話するまでである。

総発話回数 (M1): 被験者の発話回数。

総発話認識成功回数 (M2): M1 のうち、認識に成功した発話回数。

暗黙的な発話認識成功回数 (M3): M2 のうち、暗黙的な発話の認識に成功した回数。

総発話認識失敗率 (M4): M1 に対する認識に失敗した発話回数の割合。

操作完了時間 (M5): タスクが完了するまでの時間。

フィードバック削減文字数 (M6): K3 の適用によって削減された音声フィードバックの文字数。

(A) では M1, M2, M3, M4, M5, M6 を計測し、(B) では M1, M2, M4, M5 の計測を行った。M1 の値が大きいほど、M5 の値も大きくなる。また M4 が大きいほど被験者とシステムとの対話が成り立たなかったと言える。

4.5 考察

提案システムを用いた実験では、全ての被験者が理想状態の実現を達成できた。具体的な機器操作コマンドを一切伝えていないにも関わらず、適切に音声インターフェースを利用することができたという点で混合主導型対話を実現した提案システムは有効であるといえる。

次に M2 に着目すると、(A) の 6 名中 3 名が 9~11 回と (B) よりも少ない発話回数でタスクを完了している。これは、暗黙的なユーザ要求の内容によっては複数の機器が効率よく同時に制御されるためであると考えられる。一方で M1, M4 に着目すると、(A) のほうが誤認識率が高く、総発話回数が多くなる傾向が見られる。これは (A) における対話シーケンスが (B) のも

のと比較すると複雑になることがその一因であると考えられる。

特に誤認識が多かった (A) の被験者 1, 3, 5, 6 および (B) の被験者 2, 5 では、音声インターフェースによるフィードバックが終了する前に被験者が発話を行ってしまい、誤認識が発生するケースが共通して見られた。ユーザと音声インターフェースの発話タイミングの適切な制御については今後の課題としたい。

家電状態に応じたコマンドのフィードバックについては、音声インターフェースが提示するフィードバックの内容を家電状態に応じて削減することが可能となった。M6 によると平均して 29 文字が音声フィードバックから削減されたことになる。先ほど課題として述べたように、ユーザと音声インターフェース間の対話を効率化することは、誤認識の削減にもつながる可能性がある。今後は家電状態だけでなく、温度や照度といった環境要因を考慮した適切な絞り込みなども考えていきたい。

5. おわりに

本稿では、2.4 節において、現状の音声による HNS 操作インターフェースの問題点として、P1, P2, P3 の 3 つを挙げた。これら 3 つの問題点を解消するため、3.1 節において K1, K2, K3 の 3 つのキーアイデアを提案し、3.5 節でキーアイデアを実現する暗黙的なユーザ要求を抽出・推定する対話型音声インターフェースを実装した。また、4. 章では提案インターフェースの有効性を評価するための実験を行い、4.4 節で示した実験結果の考察を 4.5 節で行った。今後の研究課題として、ユーザと音声インターフェース間の発話タイミングの制御や、さらなる誤動作の低減などが挙げられる。家電機器を操作する音声インターフェースは未だ一般的なものとなっていないが、これらの課題を解決することで広く一般に利用されるシステムにより近づくことができると考えられる。

謝辞 この研究の一部は、科学技術研究費(若手研究 B 20700027, 21700077)の助成を受けて行われている。

文 献

- [1] 森川博之, “ワイヤレスが開くコピキタスネットワーク,” 電子情報通信学会誌, vol.87, no.5, pp.356-361, 2004.
- [2] M. Nakamura, A. Tanaka, H. Igaki, H. Tamada, and K. Matsumoto, “Constructing home network systems and integrated services using legacy home appliances and web services,” International Journal of Web Services Research, vol.5, no.1, pp.82-98, Jan. 2008.
- [3] 共用品推進機構, “不便さ調査データベース,” http://www.kyoyohin.org/02_syougai/index.php, 03 2008.
- [4] 渡辺裕太, 関口芳廣, 鈴木良弥, “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.
- [5] 河原達也, 荒木雅弘, 音声対話システム, オーム社, 2006.
- [6] 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.
- [7] 伊藤英一, “頭の向きと音声により操作する四肢まひ者向け家電コントローラ,” ヒューマンインタフェース学会論文誌, vol.4, no.1, pp.25-31, 2002.
- [8] 結城 浩, Java 言語で学ぶデザインパターン入門, ソフトバンククリエイティブ株式会社, 2004.