

複数家電の連携操作が可能な家電音声入力インタフェースの構築

山下浩司[†] 小木理紗子[†] 小暮悟^{††} 小西達裕^{††} 伊東幸宏^{†††}

[†]静岡学大学院情報学研究所 ^{††}静岡大学情報学部 ^{†††}静岡大学創造科学技術大学院
〒432-8011 静岡県浜松市中区城北 3-5-1

E-mail: [†]{gs07050, gs08014}@s.inf.shizuoka.ac.jp, ^{††}{kogure, konishi}@inf.shizuoka.ac.jp,
^{†††}itoh@inf.shizuoka.ac.jp

あらまし 近年、家電製品の高機能化に伴い操作方法も複雑化し、操作を覚えることが困難になってきている。複数の機器を同時に扱う場合を考えると、それはより顕著になる。そして慣れによってこの問題が解決されても、買い替え等によって使用する機器が変わった場合には、再度同じ問題に対処しなければならない。我々はこれまでに、このような問題に対処するために、ある家電機器における様々な機種を、メニュー操作も考慮した上ですべて同じ感覚で操作可能とするインタフェースをすでに構築している。しかしながら、あくまで単体の家電機器を対象としたものであったため、本稿ではさらにこれを拡張し、複数機器を対象とした連携操作を可能とした。実際に、テレビとHDDレコーダを例にとり、複数家電の連携操作が可能なインタフェースが動作することを確かめた。

Construction of speech input interface for cooperative operation of multiple home appliances

Koji YAMASHITA[†] Risako OGI[†] Satoru KOGURE^{††} Tatsuhiro KONISHI^{††} Yukihiro ITOH^{†††}

[†]Graduate school of informatics, Shizuoka University

^{††}Faculty of informatics, Shizuoka University

^{†††}Graduate school of science and technology, Shizuoka University
3-5-1 Johoku, Nakaku, Hamamatsu, Shizuoka, 432-8011 Japan

E-mail: [†]{gs07050, gs08014}@s.inf.shizuoka.ac.jp, ^{††}{kogure, konishi}@inf.shizuoka.ac.jp,
^{†††}itoh@inf.shizuoka.ac.jp

Abstract Recently, the performance of home appliances becomes higher. Accordingly it becomes difficult to learn the complicated operation methods. The problem becomes more remarkable, when we operate multiple home appliances together. Furthermore, there is another problem; that is, users have to learn the complicated operation methods again when they have new appliances. In order to solve these problems, we constructed, in the previous research, the interface by which users can operate various models of a certain type of appliance with the same operation methods. In this paper, we describe extension of this interface for cooperative operation of multiple home appliances.

1. はじめに

近年、家電製品は高機能化しており、それに伴ってその操作方法も複雑化している。そのためユーザは操作方法を覚えるのに時間がかかるようになった。また新しい機器を買い足したり買い換えたりした場合に、同じ家電であっても機種が違えば操作方法を覚え直さなければならない。このような問題に対処する方法として、学習リモコンの存在があり、一般的に利用されている。また、リモコン操作で行う機器操作をPDA等の端末で統一的に操作可能なインタフェースに関する研究[1]などもある。我々も「操作の簡略化」「機種ごとに違う操作を統一化する」という点を考慮した家電用の音声操作インタフェースを構築している[2][3]。リモコンで複数回のボタン入力が必要な操作を1回のコマンド入力力で操作可能にし（「操作の簡略化」）、同一

家電機器の複数機種を同感覚で操作可能なインタフェース（「機種ごとに違う操作を統一化する」）を構築し、実際にビデオデッキ4機種の実機を同感覚で操作可能なインタフェースを実現した[2]。そして、さらにこれを拡張し、近年の家電製品に多く見られるメニューを利用した操作に対しても、機器ごとのメニュー構造を考慮し、どの機種であっても同じ方法で同じ種類の操作が実行できるインタフェースを構築し、実際にHDDレコーダ2機種で操作を実現した[3]。本研究ではユーザからの入力に主に音声インタフェースを扱っているが、音声インタフェースには、手がふさがっている状態でも使える、訓練の必要がないなどの利点がある[4][5]。また、家電で音声インタフェースを利用する研究も行われている[6][7]。

今までの研究では、家電操作のインタフェースとい

うのは、操作対象となるものは機器単体であり、複数の機器に対して全体的な操作を行うことが出来るインタフェースは少なかった。複数機器を対象とするものの代表例としては、ホームネットワークシステム[8]やCEC規格に基づいたHDMIリンク[9]などがあるが、これらは同じメーカーや狭く限られた機種でしか利用できないことが多く、既存環境への導入や、そこからの移行を考慮しているとは言い難い。

本稿では、先行システム[3]をさらに拡張し、複数家電を1つのインタフェースで利用できるようにした。また、ユーザが操作をする際に操作対象となる機器を自動で判断し、また、ユーザが直接指示したのとは別に必要と思われる操作をシステムが自動で判断し実行する機能を実装した。

2. 機種非依存の家電音声インタフェース[3]

先行研究[3]では、「操作の簡略化」「機種ごとに違う操作を統一化する」「機器の種類や機種ごとの発話感覚を統一化する」という3つの基本方針に基づいて音声インタフェースを構築した。1つ目については、例えばHDDレコーダで保存した番組を削除する場合に、番組を選択してメニュー内をカーソル移動して削除の項目を選択する必要があるような状況で、それを一度の入力で実行できるようにするものである。2つ目については、例えばビデオデッキのスロー再生は、ある機種ではスロー再生ボタンを押すが、ある機種では、再生中に一時停止を押しその後で再生ボタンを長押しする必要があるような状況で、それを同じ操作でできるようにするというものである。先行研究では、これら1つ目と2つ目を実現するために「標準モデル」を利用している。

また、システムでは音声入力を用いているが、コマンドごとに発話の文体を覚えるのは非効率である。そこで、前述の3つ目を実現するために、入力できる文体を「統一文体」にした音声入力インタフェースを提案した。

2.1. 標準モデル

標準モデルとは、家電の種類ごとに、その家電で標準的に実行可能な機能を実現するためにどんな状態を必要とするか、どんなコマンドがそれらの状態の遷移のトリガとなりうるか、その家電がどんなパラメータを持ちえるかを定義したものであり、以下に示す(1)式で定義される。

$$M = (S, C, \delta, PM) \dots (1)$$

$$S = S_s \cup S_m$$

$$C = C_s \cup C_p$$

$$C_s = CI \cup SCI \cup MC \cup RC$$

状態セットSは機器ごとに定義した状態の集合であり、メニューに関係のない通常の状態セット S_s と、メニューに関する状態セット S_m からなる。ここでいう

状態とはビデオデッキにおける再生状態や停止状態等の機器の動作の状態であり、パラメータの変更(例えば、テレビのチャンネルやエアコンの設定温度など)などは含まない。機器が持つパラメータの情報はパラメータセットPMにまとめられる。操作セットCはその機器で利用可能な標準的操作の集合であり、状態を変化させる操作セット C_s と、パラメータの変更を行う操作セット C_p からなる。操作セット C_s は四種類の操作セットで構成される。CIは機器の基本的操作セット、SCIはリモコンのボタンを複数押さないと出来ない操作をCIに含まれる基本操作の組み合わせで実現したショートカット操作のセット、MCはメニューに関する操作セット、RCはメニュー状態におけるリモコンを用いた移動選択(上下左右、決定)に関する操作セットである。また操作セット C_s と状態セットSの組み合わせによって次状態がどのように遷移するかを表したものが状態遷移 δ である。先行研究では、SCIを定義することにより「操作の簡略化」を図っている。リモコンのボタン操作のみでは複数個のボタンを押さないといけないところを、SCIに定義されているショートカット操作セットを利用することで操作回数を削減できる。

システムは、ユーザ発話から求める操作 c (および、パラメータ値 pv)と、標準モデルM、現状態 s の3つを用いて、次状態 s' とコマンドIDの列であるCMID列を決定する。 $c \in CI \cup MC \cup RC$ である場合はCMID_s列 $\{c\}$ を決定する。 $c \in SCI$ であれば、対応する sci に登録されている ci 列を標準モデルから参照し、CMID列 $\{c_1, \dots, c_n\}$ を決定する。また、 $c \in C_p$ の場合は、CMID_p列を $\{c, pv\}$ と決定する。ここで生成されたCMID列は、その後システムの赤外線信号生成部によって実際の機器ごとの信号に変換される(これにより、機器ごとに違う操作の統一化が実現されている)。この部分の詳細は、4.2節で改めて示す。

2.2. 統一文体

ユーザが、標準モデルMで定義した操作 $c \in C_s \cup C_p$ を実際にシステムに伝える際に、その操作を音声言語でどのように表現するかを決める必要がある。例えば、全く制限のない自由発話文での入力を希望するかもしれないし、単語のみで入力する方法も考えられる。前者はユーザにとっては有効な入力手法であるが、音声認識や言語理解を考慮した場合、自由発話での入力は非現実的である。また、後者は、例えばカーナビゲーションではすでに実現されており、実現可能性は非常に高いが、多様な機能を持つ家電に対する操作を1単語ですべて表現することは難しい。そこで、本研究では、以下に示すような統一文体[10]を音声で入力するインタフェースを採用している。

対象 を 操作

属性 は 値

使用例： エアコンをつける 温度は 28度
目的地を設定する 場所は静岡県

このように文体を限定することにより、ユーザはこの二種類の文体の使い方を理解するだけで、すべての機種を同じ感覚で操作することが可能となる。また、入力できる文体をイメージしやすくなるという利点もある。

現象を表すような表現は、すべて「属性は属性値」という形に同義変形できることがわかっている[11]。ただし、例えば「状態は再生」のような表現は、不自然なため、「対象を操作」という形式(例えば「テープを再生」)も発話可能にしている[10]。

操作セット C の要素である操作 c のそれぞれには統一文体を用いた発話が割り当てられ、複雑な構文解析、意味解析をすることなくユーザの発話を該当する操作 c に単純に変換可能である。

3. 複数操作の連携操作が可能な家電用音声インタフェース

本研究の狙いは、単体家電を対象としていた先行システムを拡張して複数家電の操作に対応させることでユーザの複数家電操作における操作の煩雑性を解消し、システムを用いた機器操作の利便性をより向上させることである。

操作対象がひとつではなくすることで、システムは最初に操作対象となる機器を判定する必要がある。また、先行システムでリモコン操作による操作ステップを簡略化したように、複数の機器への機械的な操作をまとめて行えるようにすることで操作ステップの簡略化を行い、システムの利便性向上を図る。

複数家電を操作対象とする場合、標準モデルは家電の数だけ用意され、システムはその内、ユーザが操作を行った家電の標準モデルを利用し動作する。標準モデルがどの家電のものか識別できるように、標準モデル M に拡張を行い、拡張標準モデル M' を以下のように定義する。ここで、 Hn_i は、家電を識別するための名前である。

$$M'_i = (S_i, C_i, \delta_i, PM_i, Hn_i)$$

3.1. 対象機器の推定

複数機器を操作対象とする場合、システムはユーザからの要求がどの機器に対するものなのかということを最初に判断しなければならない。システムが操作対象となる機器を判断するには次のパターンを考える必要がある。

- (i) ユーザによって明示的に対象機器の指定がされる場合
- (ii) ユーザによって明示的に対象機器の指定がされなかった場合

(i)の場合に対処するため、「(対象機器の指定) + 統一文体」の発話でも受理できるように音声インタフェースの拡張を行った。対象機器の指定には、家電の種類 Hn_i (テレビや HDD レコーダなど) のほかに、機器名自体の名前 DN_{Hn_i} (ブラビアやレグザ, RD-X6 など) でも利用可能とした。

(ii)の場合には、操作する対象機器を推定する必要がある。つまり、以下の場合を考慮する必要がある。

- (a) 推定された対象機器が 0 個
- (b) 推定された対象機器が 1 個
- (c) 推定された対象機器が 2 個以上

まず、(a)の場合であるが、本研究における音声インタフェースで受け付ける入力文は、必ずどれかの機器の操作に対応づけられるため、考慮する必要がない。次に、(b)の場合は、推定された対象機器について操作を行えばよい。最後に(c)の場合であるが、この場合は、システムでは機器をひとつに推定できなかったため、ユーザに対し機器名を聞く応答をして、ユーザに機器を指定してもらえばよい。

対象機器の推定の仕方であるが、例えば、電源を消す機能を持つ家電 A,B,C が操作対象となっており、ユーザから「電源を消す」という発話を受理したとき、家電 A,B の電源が既に消えている状態であれば、発話の中で対象を家電 C と指定していなくてもユーザの発話が家電 C に対するものだと判定できるので、対象機器は家電 C だと推定し、家電 C に対して操作を行えばよい。

3.2. 複数家電の連携操作

一般的に、家庭内には多種類の家電があり、ユーザはそれらを随時操作している。このとき、ひとつの機器だけに操作を行う場合と複数の機器に対して操作を行う場合がある。複数の機器を連続して操作する状況は、以下の二種類に分けられる。

- (a) 対象の機器が連動するように接続されており、ユーザが希望する状態を実現するために、それら連動する機器の各々を操作する必要がある場合
- (b) 複数の独立した個別要求を一度に満たそうとした場合

(a)は「アンプをつけてから DVD を再生する」というような場合である。このような場合に、ある機器が特定の状態の時に、接続された他の機器がどのような状態かという情報をルールにしてシステムに設定しておく。そして(a)の状況が発生したときに、あるひとつの機器への特定の操作をトリガとして、ルールを満たす一連の操作を自動で実行するようにするのが本研究という連携操作である。なお、(b)についてはユーザの習慣や利用環境に大きく依存し、特定の操作を一般化するのが難しいため、本研究では対象としない。

連携操作は、以下に示す連携操作セット COP にて定義される。

$$\begin{aligned} COP &= \{COP_1, \dots, COP_n\} \\ COP_i &= \{Hn_{ij}, S_i, PRES_i\} \\ PRES_i &= \{PRE_{i1}, \dots, PRE_{in}\} \\ PRE_{ij} &= \{Hn_{ij}, S_{ij}, PM_{ij}, PV_{ij}\} \end{aligned}$$

ある連携操作 COP_i は、家電 Hn_i が状態 S_i になるとき、 $PRES_i$ で指定された家電 Hn_{ij} が状態 S_{ij} になって

いる（もしくは、パラメータ PM_{ij} が値 PV_{ij} になっている）必要があるという情報を記述している。こうすることにより、ある家電である操作を行う前に、ほかの家電をふさわしい状態（あるいはふさわしいパラメータ）に変更することが可能である。

例えば、 $COP_1 = \{\text{HDD レコーダ, 再生状態, \{\{テレビ, 電源オン, null, null\}, \{テレビ, null, 入力, 外部入力2\}\}}\}$ のように定義されている場合、HDD レコーダに対して再生を実行する際、まず、テレビの電源が入っているかどうかを確認し、入っていなかったら、テレビの電源を入れるコマンドを発行し、そののち、テレビの入力が外部入力2になっているか確認し、なっていないければ入力切り替えをするコマンドを必要回数だけ発行した後に、HDD レコーダに対して再生のコマンドが発行されることとなる。

4. システム構成

4.1. 入力処理部

入力処理部はユーザの発話を受理し、発話内容、各機器の状態などの情報を元に、機器に行うべき適切な操作を示すコマンド ID 列を決定する。入力処理部は操作対象となる機器情報 H_{ni} と行うべき操作を示すコマンド ID 列（CMID 列）を、リモコン信号生成部に送信する。図 1 に入力処理部の概要を示す。

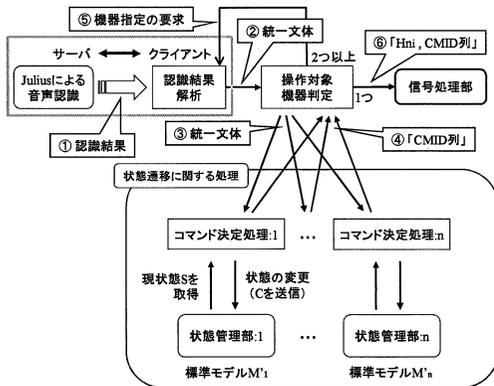


図 1: 入力処理部の概要

4.1.1. 操作コマンドの決定

音声認識にはオープンソースの音声認識システムである Julius を利用する[12]。まず、Julius を音声認識サーバとして起動しておく。入力処理部では、音声入力を受け付ける段階で、Julius にクライアントとして接続し、認識結果を受け取る（図 1 ①）。受け取った音声認識結果を統一文体の文体（図 1 ②、③）を考慮して解析し、目的の操作（図 1 ④）を決定する。自然発話文による音声認識結果を利用する場合は、構文解析、意味解析等を行った後に、どの操作に対応するかを理解しなければならないが、統一文体を利用する事によって、複雑な構文解析等をする必要はなく、曖昧

性なく操作を決定することが可能である。

機器の操作 c は機器の状態を変化させるものと、機器のパラメータを変更するものの二種類に分けられる。ユーザの発話が前者だった場合、対象となる家電 H_n の現状態 s_i を確認し、標準モデル M' を参照する。そして、現状態 s_i と発話から求まる操作 c_s を元に状態遷移 δ を参照し、その家電が次に遷移すべき次状態 S_j を決定する。後者であった場合は、発話内容から、パラメータ変更を行う操作 c_p とパラメータの変更内容 pv （絶対値指定か相対値指定）を判断する。これらは図 1 中のコマンド決定処理と状態管理部によって判定、処理される。

4.1.2. 操作対象の決定

先行システム[3]ではユーザの入力は事前に決められたひとつの機器に対するものであるため、常に同じ標準モデル M を利用して動作していた。しかし複数家電を操作対象とする場合、利用される標準モデル（拡張した標準モデル M' を用いる）は複数になる。図 1 では n 個の標準モデルを利用して n 個のプロセス（コマンド決定処理、状態管理部）が動作している。そのため、発話内容からどのモデルを利用し、どの機器に対して操作を行うかということを判定する必要が生じ、これを処理するのが図 1 の操作対象機器判定である。

システムは、対象となる機器を識別する必要があるが、ユーザによって明示的に機器名が指定されるかどうか、すなわち 3.1 節で示したパターン(i),(ii)の二種類を考慮する必要がある。ユーザが発話の中で対象機器を明確に指定する場合、すなわちパターン(i)であった場合は、ユーザの指定した対象機器の標準モデル M' を利用して操作 c の決定を行う。ユーザが発話の中で対象機器を指定しなかった場合、システムはユーザの発話を受理すると、最初に、読み込んでいるすべての標準モデル M' を対象として先行システムと同様のプロセスで解析を行い、ユーザの発話に該当する操作 c を探す。このときひとつの操作 c を決定できた場合（パターン(ii)(b)の場合）は、リモコン信号生成部に対して「機器名 H_{ni} + 操作 CMID 列」の形で情報を送信する（図 1 ⑥）。もしひとつに決定できなかった場合（パターン(ii)(c)の場合）は、ユーザに対して家電の種類 H_{ni} 、あるいは機器名 DN_{ij} によって対象となる機器を明示的に指定するように要求する（図 1 ⑤）。

4.2. リモコン信号生成部

リモコン信号生成部は、入力処理部から受け取った操作コマンド ID（CMID_s もしくは CMID_p）を実際に機器に送信する信号の形に変換し、送信をする部分であり、家電機器に依存する部分である。

まず、機器に付属しているリモコンのボタン一つ一つに ID を割り振る。この ID をコード ID と定義する。このコード ID と実際の赤外線信号 IR を 1 対 1 で CODE_IR として定義する。

$$CODE_IR = \{\{code_1, ir_1\}, \dots, \{code_n, ir_n\}\}$$

これに対し、入力処理部が出力する家電種類ごとに定義された(すなわち、家電機器自体には依存しない)コマンド ID 列とコード ID の対応を COMMAND_CODE として定義する。

$$\text{COMMAND_CODE} = \left\{ \begin{array}{l} \{\text{CMID}_1, \text{code}_{1,1}, \dots, \text{code}_{1,m_1}\}, \\ \dots \\ \{\text{CMID}_n, \text{code}_{n,1}, \dots, \text{code}_{n,m_n}\} \end{array} \right\}$$

メニュー構造リスト MENU_STR は機器ごとのメニュー構造を記述したもので、機器のメニュー項目 im がどのように構造化されているか定義されている。MENU_STR は次のように定義される。

$$\text{MENU_STR} = \{\text{im}_1, \dots, \text{im}_{n_{im}}\}$$

im にはそれぞれメニュー ID が振られており、これは MC と対応している。入力処理部から MC を受け取ると対応した im を MENU_STR から探し、実行するのに必要な操作手順をコード ID 列として生成する。

本研究では複数の機器を操作対象としているため、信号生成部は操作対象となる機器それぞれに対応した COMMAND_CODE, CODE_IR, MENU_STR を使用する。

信号生成部は、入力処理部から操作対象となる機器の情報 H_{m_i} と、コマンド ID (CMID_s もしくは CMID_p) の列を受け取る。そして対象となる機器の COMMAND_CODE もしくは MENU_STR を参照し、対応する code および im 列を決定する。最後に、CODE_IR を参照して、ir 列を決定し、赤外線信号として送信する。

ここで、コマンド ID をコード ID の形に一度変換している理由は、機種依存の操作方法を吸収するためである。これは、コマンド ID が同じでも送信しなければいけないボタンの信号は違う場合があるからである。このような操作例としては、2 章で挙げたスロー再生などがある。ビデオデッキのスロー再生では、あるメーカーはスロー再生ボタンがあるが、あるメーカーにはスロー再生ボタンはなく、再生中に一時停止ボタンを押したあと、再生ボタンを長押しする必要がある。最後に、生成されたリモコン信号は赤外線デバイスを通して機器に送信されることで、機器が指定された操作を実際に実行する。

5. テレビと HDD レコーダへの適用

複数機器への操作を行う例として液晶テレビと HDD レコーダを選択した。AV 機器は一般に複数の機器を同時に動かすという状況が多く、近年の高機能化によって操作が難しくなっているもののひとつである。そのような問題の解決法のひとつとして 1 章で述べた HDMI リンクがあるが、これにはメーカーや対応機種などの制約がある。本研究では機種の違いによる制約は受けたくない実装を行っており、実際にシャープ製液晶テレビ LC-16E5 と東芝製 HDD レコーダ RD-X6 に対してシステムを適用した。これらの家電はメーカーが異なり、

また RD-X6 は HDMI リンク普及前の製品であり HDMI リンク非対応製品である。

液晶テレビと HDD レコーダの標準モデル M' における S_s, S_m, C_p, CI, SCI, MC, RC の個数を表 1 に示す。また、実際に適用した LC-16E5 と RD-X6 それぞれについて、機種依存情報である COMMAND_CODE, CODE_ID, MENU_STR の個数を表 2 に示す。

表 1：標準モデルについて

	S _s	S _m	C _p	CI	SCI	MC	RC
液晶テレビ	7	0	16	12	5	0	0
HDD レコーダ	42	4	4	47	61	24	5

表 2：機種依存情報について

	LC-16E5	RD-X6
COMMAND_CODE	35	88
CODE_ID	59	57
MENU_STR	0	89

テレビと HDD レコーダの連携については、HDD レコーダの S_s42 状態のうち、40 状態でテレビとの連携ルールが設定されている。テレビはテレビ単体で行う操作しかないので、テレビの S_s にはルール設定されたものはない。連携ルールが設定された状態 s_s に遷移する操作 c が選択されると、ルールに設定された他の家電が設定された状態になるように操作が行われる。

このような状態で、実際に「HDD の再生を行う」という操作を行った場合について、(i)リモコンを用いた場合、(ii)先行システムを利用した場合、(iii)本システムを用いた場合の操作手順について以下に示す。

対象家電はテレビ、HDD レコーダとし、ともに電源 OFF の状態であるとする。また、HDD レコーダはテレビの外部入力 2 に接続されており、テレビの入力切り替えは TV の状態であるものとする (つまり、外部入力 2 に切り替えるためには、入力切り替えが 2 回必要であるとする)。

(i) リモコンの場合

- ① テレビのリモコンの[電源ボタン]を押し、テレビの電源を入れる
- ② テレビのリモコンの[外部入力ボタン]を 2 回押し、テレビの外部入力を HDD レコーダ(外部入力 2)にする
- ③ HDD レコーダのリモコンの[電源ボタン]を押し、HDD レコーダの電源を入れる
- ④ HDD レコーダのリモコンの[再生ボタン]を押し、HDD レコーダで再生操作を行う

(ii) 先行システムの場合

- ① テレビ用の音声インタフェースに「電源を入れる」と発話し、テレビの電源を入れる。
- ② テレビ用の音声インタフェースに「入力外部入力 2」と発話し、テレビの外部入力を HDD レコーダにする。

③ HDD レコーダ用の音声インタフェースに「HDD を再生」と発話し、HDD レコーダで再生を行う。
(iii)本システムの場合

① 統合音声インタフェースに対して「HDD を再生」と発話することで、HDD レコーダで再生操作を行う。

(ii) における発話回数は 3 回で、(i) における操作回数 5 回に比べ、操作手順が減っている。これは、HDD レコーダの音声インタフェースにおいて、電源が入っていない状態から HDD を再生するには電源を入れてから再生をする必要があるという情報を標準モデル M が持っていることで実現している。また、②の外部入力 2 という発話に対しても、現在の入力がどこになっているかを判断し、自動で適切な回数の入力切り替え信号を送信することで実現している。

(iii) における発話回数は 1 回で、(i)の操作が一度の発話で行えるようになっており、(ii)に比べても同様である。また、(ii)ではシステム利用時に音声インタフェースを切り替える必要があったが、(iii)では対象の指定なしに操作をしている。このように本システムを利用することで、ユーザによる機器の操作回数を減らし、操作の簡略化が可能であることが確認できた。

6. まとめ

異なる機種に対し同じ感覚で操作を行うことが出来る単体家電用の音声インタフェースを拡張し、複数家電の連携操作を可能にした家電音声インタフェースを構築した。この音声インタフェースは、ユーザが対象機器を直接指定しなくても各機器の状態からシステムが対象を推測できる。またある家電のある操作に対して、別の家電の操作が別途必要である場合に、システムが自動で別の家電を操作する。これにより、ユーザによる機器操作の手順の削減を図った。液晶テレビ LC-16E5 と HDD レコーダ RD-X6 に対して実際に適用し、リモコンを用いた操作と先行システムを利用した操作よりも、操作手順を減らすことが確認できた。

今後は、音声インタフェースを実際に被験者に使ってもらって客観的評価実験を行う予定である。また、個人による機器操作はユーザの個人的な習慣や、機器の置かれた環境などによって、完全な一般化はされなくとも、ある程度固定化されることも多いと考えられる。そのため、ユーザの行動履歴などを考慮した上で機器の推定を行うことが出来るシステムを構築できればユーザによる機器選択の手間が減る。また、3.2 節の(c)のうち、特定の条件下で複数の機器に一定の操作を行うことが生活習慣となっている場合、例えば、「出掛けるときに、照明とテレビを消す」といったものがある。このような場合についてもユーザが「出掛ける」という入力をした時に、照明とテレビを消すと言うような連携操作にも対応していきたい。このような改良を加えることで、システムによる複数機器への操作利便性はさらに向上するのではないかとと思われる。

参考文献

- [1] Jeffrey Nichols, Brad A. Myers, Michael Higgins, Joseph Hughes, Thomas K. Harris, Roni Rosenfeld, Mathilde Pignol, "Generating Remote Control Interfaces for Complex Appliances", Proc. of the 15th Annual ACM Symposium on User Interface Software and Technology UIST'2002, ACM Press (2002).
- [2] 小川耕司, 小暮悟, 小西達裕, 伊東幸宏, "異機種を統一的に利用可能な家電統合インタフェースの構築", 情報学ワークショップ 2007 (WiNF2007) 論文集, pp.99-102 (2007).
- [3] 小川耕司, 山下浩司, 小暮悟, 小西達裕, 伊東幸宏, "メニュー操作を考慮した機種非依存の家電統合インタフェース", 情報処理学会研究報告 SLP-70, pp.89-94 (2008).
- [4] 甲斐彦彦, 盛浩和, 中野崇弘, 中川聖一, "フォーム型 Web 情報検索サービスのための音声ユーザインタフェースシステムと操作性の評価", 情報処理学会論文誌, Vol.46, No.5, pp.1318-1329 (2005).
- [5] 原直, 白勢彩子, 宮島千代美, 伊藤克互, 武田一哉, "音声対話による楽曲検索システム", 情報処理学会研究報告 SLP-53-6, pp.31-35 (2004).
- [6] 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).
- [7] 吉田諒, 吉田恵, 安村通見, "音声とテンキーを用いた視覚障害者向けリモコンの試作と評価", ヒューマンインタフェース学会論文誌, Vol.9, No.2, pp.87-92 (2007).
- [8] 榊原弘記, 中村匡秀, 井垣宏, 松本健一, "ホームネットワークシステムにおける家電状態を利用した音声操作インタフェースの改善", 電子情報通信学会 2 種研究会サイバーワールド (CW) 第 9 回研究会, pp.13-18 (2008).
- [9] HDMI を利用したシステムコントロール (HDMI-CEC: High Definition Multimedia Interface-Consumer Electronics Control) HDMI Licensing <http://www.hdmi.org>
- [10] 森田浩康, 桂川桂子, 林真弘, 伊藤敏彦, 甲斐彦彦, 小西達裕, 伊東幸宏, "車載器との音声言語インタフェースにおける文体統一の効果", 情報学ワークショップ 2003 (WiNF2003) 論文集, pp.185-189 (2003).
- [11] 池ヶ谷有希, 野口靖浩, 小暮悟, 伊藤敏彦, 小西達裕, 近藤真, 麻生英樹, 高木朗, 伊東幸宏, "対話文脈を利用した構文意味解析", 人工知能学会論文誌, Vol.22, No.3, pp.291-310 (2007).
- [12] 大語彙連続音声認識エンジン Julius <http://julius.sourceforge.jp>