

親和行動導入による実用的ホームロボットインタフェースの研究 — ユーザの言葉で操作・通知するインタフェース —

山本 大介^{*1} 土井 美和子^{*1} 松日楽 信人^{*1} 上田 博唯^{*2} 木戸出 正継^{*3}

^{*1}(株)東芝 研究開発センター ^{*2}(独)情報通信研究機構 ^{*3}奈良先端科学技術大学院大学

概要—情報家電の普及に伴い、ユーザが音声で自然に操作できるインタフェースが望まれている。しかし現状、音声の誤認識は避けられず、また決った指示語以外は受け付けられない。本研究では、間違っても許され次第に言葉を覚える幼児を参考にした親和性をロボットに導入しこの解決を図っている。本稿では、誤認識が許されるうちに、幼児が言葉を覚えるようにロボットがユーザの音声指示語を獲得する方法とその評価について報告する。

Practical Robotic Interface Innovated by Familiar Behaviors - Interface by User-defined Instructions -

Daisuke Yamamoto^{*1}, Miwako Doi^{*1}, Nobuto Matsuhira^{*1}, Hirotada Ueda^{*2}
and Masatsugu Kidode^{*3}

^{*1} Corporate Research & Development Center, Toshiba Corporation

^{*2} National Institute of Information and Communications Technology

^{*3}Nara Institute of Science and Technology

Abstract - This paper proposes a practical robotic interface that operate by user-defined instructions. Recently, a robotic interface attracts people's attention as an intermediary between complicated information home appliance and its user. The user wants to operate the appliances by voice. There are two problems. The one is that the voice recognition is difficult especially at home and by inexperienced users. The other is that the voice instructions are decided beforehand. We think if the robot acts like an infant, these problems are solved. The user will accept recognition error by the robot. And the user will be able to teach the user-defined instructions, naturally. Based on this idea, we innovate infant's familiar behaviors to the robotic interface. This paper presents the effects of the familiar behaviors.

Keywords : Robotic Interface, Voice Recognition, Infant's Familiar Behaviors

1. はじめに

近年、ブロードバンドの普及に伴い家庭内でも情報家電をはじめとするホームネットワークの構築が進んでいる。さらに、これらの情報家電と各種センサが融合したユビキタスネットワークに関する研究も盛んである。しかし、これら情報家電のユーザインタフェースは、必ずしも使いやすいものとは言えない。それは各家電が、例えば、冷蔵庫の在庫を示したり、EPG(Electronic Program Guide)と連携し番組表を見ながら番組を選んだり、

と多くの便利な機能を持ち、様々な使い方ができるようになる一方、その多様さがゆえに、ユーザの欲する機能を実行するために複雑な入力が強いられるためである。この問題を解決するために、情報家電とユーザとの仲立ちとなり、簡単に操作できるインタフェースとして、音声で応答し親しみやすい身体性を持つロボット用いたインタフェースが注目されている^{[1],[2]}。しかし、これには2つの課題がある。一つはロボット側の音声認識技術の課題である。家庭環境の多様な騒音に加え、

ロボットの音声認識ではマイクと発話者の位置関係が定まらず音響特性は変動する。さらに子供から老人まで幅広くかつ不慣れなユーザに対応する必要があり、その誤認識は避けられない^[3]。もう一つは情報家電側の課題である。従来からあるリモコン等の手で操作する入力に代わり、音声入力を用いた音声リモコンの研究がされている^[4]。しかし個々の操作は容易になるが、多くの機能を持った情報家電を音声で扱うには、予め決められた音声指示語を覚えなければ使えず、ユーザには大きな負担となる(図1)。

本研究では、身体性を持つロボットインタフェースに、幼児を参考にした親和行動を導入することで、まずは音声誤認識が許容されることを目指す。次に誤認識が許されるうちに、徐々にロボットがユーザの音声指示語の獲得し、決められた言葉以外で情報家電を操作する方法を提案している。そして、これまでに幼児を参考とした3つの親和行動(発話・注視・移動動作)を実装し、28名の一般被験者での評価実験により、音声誤認識の許容させる効果を確認している^[5]。

本稿では、誤認識が許されるうちに、幼児が言葉を覚えるようにロボットがユーザの音声指示語を獲得する方法とその評価について報告する。

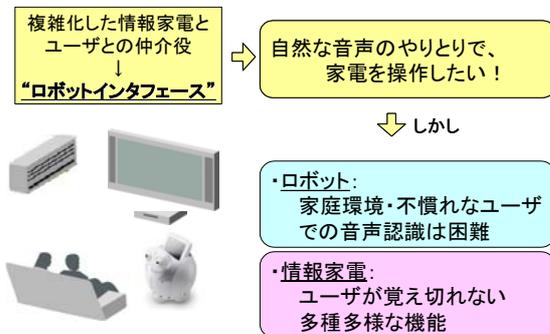


図1 ロボットインタフェースの課題

2. 親和行動の導入

前章で述べたように、ロボットインタフェースの2つの課題を解決するために、我々は、親しみを感じやすい幼児の親和行動を参考にしている(図2)。これは幼児の持つ、誰もがやさしく接したくさせる面と、人の指示を理解しようとする面を利用したいという考えからである。まず、幼児の優れた親和行動を取り入れロボットが幼児

のようにふるまい、ユーザとロボットの間、親と幼児の関係をとり入れようとしている。そして親の立場に立ったユーザは、自然と明確な分かりやすい指示を行い、親が幼児の過ちを許すように、ユーザはロボットの誤認識が許容される。また、幼児が分からないことがあると親に聞くように、ユーザが家電を操作した時に、ロボットの分からない操作があれば、ユーザに聞く。ユーザは、親が幼児に聞かれたらやさしく答えるように、ロボットにその操作の意味(言葉)を答える。結果、幼児がだんだんと親の言うことを理解しそれに適した行動をとるように、ロボットは、ユーザの教えた家電操作の意味(言葉)を覚え、その言葉に適した操作をすることができるようになることを目指している。

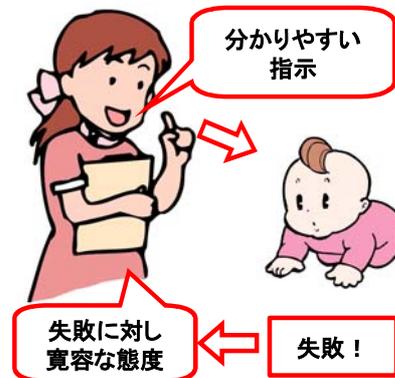


図2 幼児の親和行動の効果

こういった幼児の親和行動を実装し、評価するために、(株)東芝で開発中のロボット情報家電 ApriAlphaTM^[6](表1, 図3)を用いた。本ロボットの身体構成としては、視線を変え表情を作る“顔”である2眼のカメラ雲台と、対象物へ移動する“脚”であり、体全体の向きを変える“腰”でもある2輪独立駆動方式の移動台車からなる。また目が大きく全体に丸みを帯びた形状は、ローレンツの幼児図式(baby schema)^[7](人間の幼児や動物の幼体のもつ、丸い体型、身体に比較して大きな頭、まるまるとした手足や頬などの形態的特徴)にかなっている。これは成人・成体に「かわいらしい」とか「愛らしい」という感情や気持を起こさせると言われている。また、“顔”の特徴として、一般のロボットでは、首のパン・チルト動作により、上下左右の視野を確保するのにに対し、本ロボットでは、首のパン軸に加え眼球にパン・チルト軸があり、多くの表情を作り出している。

表1 ロボット情報家電 ApriAlpha™仕様

寸法	直径φ350mm 高さ380mm
質量	約9.5kg
ユーザ インタフェース	CCD カメラ,マイク,スピーカ 液晶モニタ(タッチパネル付)
移動速度	約0.5m/s
通信	無線 LAN(IEEE802.11b)
電源	リチウムイオンバッテリー (駆動時間 連続2時間)



図3 ロボット情報家電 ApriAlpha™概観

3. 親和行動による音声誤認識の許容⁵⁾

音声誤認識に対する親和行動の許容効果を調べるため、幼児の親和行動の中で最初に表れるコミュニケーション動作と言われる注視・発話動作、また運動系の発達に伴って現れる移動動作を実装した。以下、幼児の発達過程と対応する動作⁸⁾について述べる。注視動作として、幼児には、共同注意という母親と同じ物を見るという動作がある。ロボットでは、顔方向追従をする機能を実装している。また幼児が人の顔を好むことも知られており、ロボットにも人の顔を検出しその方向を見る動作を入れた。発話動作として、幼児は、喃語発話、音韻模倣発話を経て、一語文を発話するようになる。ロボットには一語文発話期を想定し、ユーザの発話を単語レベルで認識、その認識単語を2回繰り返して発話させた。移動動作としては、這い這い、伝い歩きを経て、興味のある方へ行く探索動作が現れる。ロボットでは、車輪による移動ではあるが、操作対象への移動動作を実装した(図4右)。

上記の音声誤認識の許容への有効性を調べるため、高齢者、若者の計28名の被験者によるSD法による印象評価実験を実施した。実験内容は、ユーザの音声指示(2チャンネル(以下ch.)にして“etc.”)によりロボットがTVのch.を切り替えるタスクを設定。親和行動としてユーザの指示を

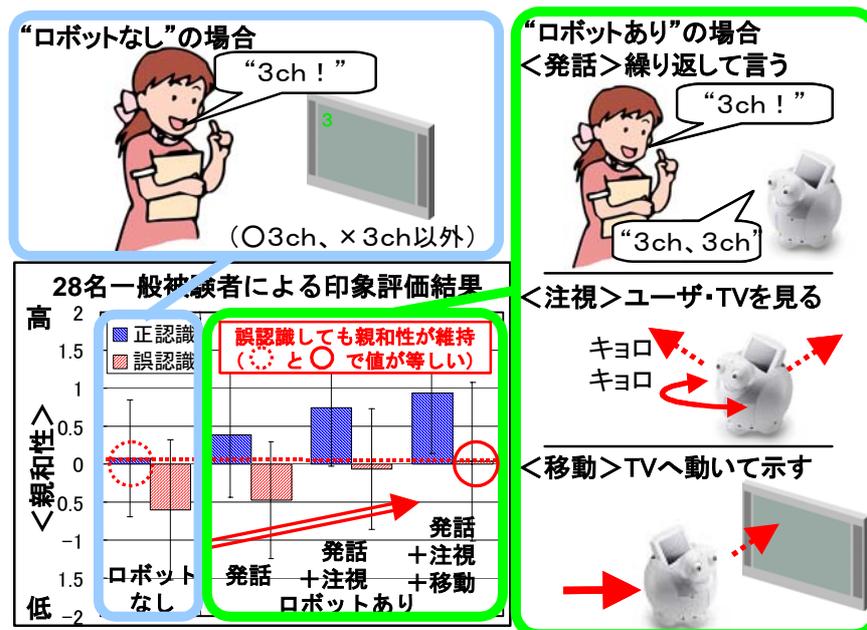


図4 誤認識許容に対する親和行動の効果

繰り返す発話動作、ユーザとTVを交互に見る注視動作、ch.を切り替える際はTVの近くへ移動する移動動作を順に加えていき、評価基準として、ロボットなしで、テレビに直接音声指示するパターンと合わせ計4パターンの比較を行う。実験は、各パターン4試行（指示が達成されるまで繰り返す）を行い、誤認識時の影響評価のため50%の誤認識を加えた。そして各パターン終了後に”冷たいー暖かい”、“暗いー明るい”などからなる7段階22形容詞対からなる質問紙に、正しく認識した場合と誤認識した場合の印象に分けて回答を得た。また形容詞対の選択には、[9]を参考にした。

印象評価の因子分析の結果、“感じの悪いー感じの良い”、“親しみにくいー親しみやすい”の順に負荷が高い第1因子（寄与率約57%）が得られ、ユーザとの親和性を示す“親和性因子”と名づけた。親和性因子は、音声認識の正誤に関わらず、発話・注視・移動の親和行動を加えていくにつれて増加し、発話+注視+移動を加えて音声で誤認識した際と、ロボットなしで音声を正しく認識した場合がほぼ同じとなった（図4左下グラフ）。このことから、ロボットの親和行動が誤認識を許容させる効果があったと言える。

また、この結果を若者（20代）と高齢者（60代）に分けて集計したところ、高齢者は若者よりも親和行動による親和性の効果が高く、特に音声を誤認識した際にもあまり影響を受けないという結果が得られ、この親和行動が、高齢者向けに対して特に有効であることが分かった。

4. 親和行動を利用したユーザの言葉の獲得

4.1 ユーザの言葉の獲得・操作

予め決められた言葉ではなく、ユーザの言葉で家電を操作するためには、ロボットが家電操作の音声指示語をユーザから得る必要がある。しかし、例えば、キーボードを用いて音声指示語を登録するといった方法は使いにくく、多くの操作コマンドを登録するのは現実的ではない。本研究では、自然な音声のやりとりでユーザの言葉を獲得する方法を考案した。これは、幼児が分からないことを親に聞くように、ロボットが分からない操作をユーザに聞き、教示してもらおうことで、ユーザの言葉と家電操作の対応付けを獲得するものである。

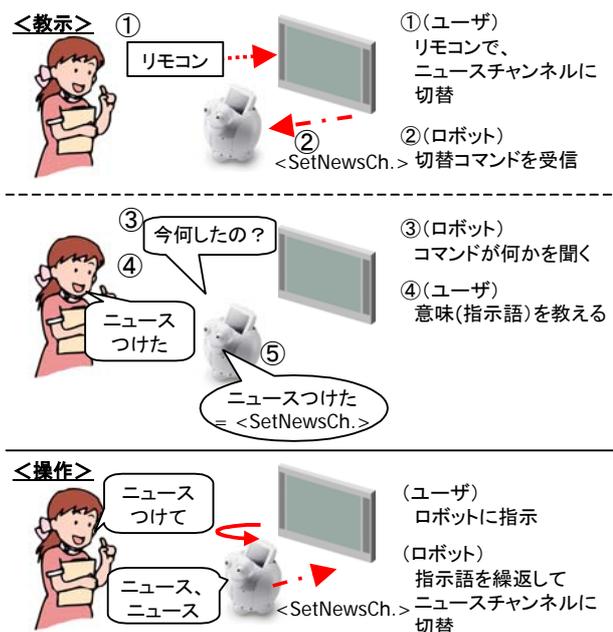


図5 ユーザの言葉の教示・操作

具体的に、多チャンネル時代のテレビを想定し、ニュースチャンネルに切り替える場合を用いて説明する（図5）。教示する場合、ユーザは、通常と同じようにリモコンを使ってテレビを操作する。その際にロボットは、ネットワークに接続された家電であればネットワークを経由して、そうでない家電であればリモコン信号を直接ロボットが受信し、家電が操作されたことを検出する。そして、ロボットはそのコマンド（ネットワーク家電であれば、例えば<SetNewsCh>、リモコン信号であればその信号コード自体）が今までに操作されたことのないコマンドで、それに関する情報を持っていないければ、“今何したの？”とユーザに質問する。それに対して、ユーザが、“ニュースつけた”と答えたら、ロボットは、認識した言葉を復唱し、“ニュースつけた”と先のコマンド<SetNewsCh>の対応を獲得する。

操作する場合は、“ニュースつけて”と指示し、ロボットは、“ニュース、ニュース”と言いながら、テレビをネットワーク経由で（もしくはリモコン信号を発信して）チャンネルを切り替える。

また教示時と操作時の語尾の違い、“ニュースつけて”、“ニュースつけた”の語尾（“で”、“た”）に関しては、文節で切り分け、“ニュース”+“つけ”+“で”と、個々に<SetNewsCh>に対応付け、操作する際にも、ユーザの発した言葉を文節で切り分け、各文節の頻度と合わせたコマンドとユー

ザの指示語の適合度を算出することで、“ニュースつけて”でも操作することを可能としている。そして、“つけ”、“で”などの頻出する語については、様々な操作を教示する際にも発されるため適合度が下がり、それらの語だけでは操作はされなくなる。

教示する際に、音声を正しく認識しなかった場合（例えば、“ニュースつけて”を“入試つけて”と誤認識した場合）、ロボットの復唱を聞いて、言い直すか、再度リモコンを用いて操作する。ロボットは、学習が進んでいなければ“今何したの？”と再度質問し、学習がある程度進んでいれば既に学習した言葉（“入試つけて”）を発話する。もし、この発話が間違っていれば、ユーザが再度、教示し、次第にロボットは正しい指示語を学習することができる（図6）。

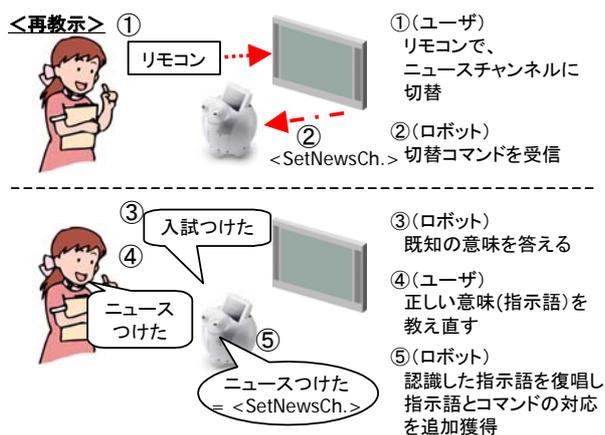


図6 ユーザの言葉の再教示

5. 実装と検証

5.1 実装

ユーザの言葉の獲得・操作方法の検証のため、4章で述べた方法を ApriAlpha™ に実装した。操作する家電としては、AV 機器の例としてテレビ、ネットワーク家電の例として UPnP™ [10] で接続されたネットワークエアコン、ネットワークに接続されてない家電の例としてリモコン操作可能な電灯スタンドを準備した。その操作としては、テレビは3つのチャンネルの切替、エアコンは冷房・暖房のオンオフ、電灯は点灯・消灯の計7操作を用意した。

5.2 検証

検証としては、初めにリモコンを用いて操作、その際のロボットの発話に応じて教示した後、適宜、音声による各家電の操作を行った。図7にその際の様子を示す。この結果、自然な対話の中で、先の7操作に対する指示語を教示でき、教示が進むと、ユーザの言葉で各家電の操作が可能になることが確認された。

さらに、先に述べた以外の2つの効果を確認した。一つは、音声誤認識に対する効果である。例えば、図5の例で、“ニュースつけた”を“入試つけた”と誤認識した場合は、“入試つけた”を <SetNewsCh> と対応付けをしてしまう。ただ操作する際にも同じように“ニュースつけて”が“入試つけて”と間違って認識される可能性が高く、誤認識も含めて対応付けをすることは、音声誤認識をしても正しく操作できる、すなわち音声誤認識に対してロバストにできる。

もう一つは、ユーザの慣れを引き出す効果である。教示・操作をする課程で、ロボットの認識結果を発話するため、ユーザがロボットの認識の癖を知り、自然とロボットの認識しやすいように発話できるようになる。また認識した発話を返すことが、簡単なユーザとロボットとの対話に繋がる利点があることも分かった。



図7 ApriAlpha™ での検証の様子

6. おわりに

情報家電への実用的なインタフェースとして、音声誤認識が生じる環境下で、音声で操作可能なロボットインタフェースを提案し、誤認識が許されるうちに、幼児が言葉を覚えるようにロボットがユーザの音声指示語を獲得する方法を提案、実装した。その結果、自然な対話を通じて、ユーザの音声指示語を獲得し、家電を操作するロボットインタフェースを実現した。さらに、音声誤認識を含めて学習することで、音声誤認識に対してロバストなインタフェースが構築できることを確認した。

今後、このロボットインタフェースを複数の一般被験者での評価実験と、実際に生活する場での評価実験を行い、定量的に評価する。そして、音声誤認識のある環境下で、自然な音声指示で操作できるで実用的なホームロボットインタフェースの構築を目指す。

謝辞

本研究を進めるにあたり、UKARI プロジェクトサービス WG で、貴重な意見を頂いている WG メンバー各位に深く感謝致します。

参考文献

- [1] 上田：ユビキタスホームにおけるサービスと対話インタフェースロボットの試作，電子情報通信学会モバイルマルチメディア通信研究会(MoMuC), Vol.105, No.264, pp.1-4(2005)
- [2] 新庄他：マルチモーダル対話技術による知的ユーザーインタフェース，日立評論，Vol.87, No.10, pp.47-50(2005)
- [3] 石川他：音声インタフェースの実用化の現状と今後の課題，情報処理学会第 66 回全国大会 (2004).
- [4] 吉田他：音声リモコン方式を用いた家電操作の試作と検討，情報処理学会第 116 回ヒューマンインタフェース研究会研究報告，2005-HI-116, pp.65-70 (2005).
- [5] Daisuke Yamamoto, et al. : Familiar Behaviors Evaluation for a Robotic Interface of Practicality and Familiarity, The 4th IEEE International Conference on Development and Learning (ICDL-05) (2005).
- [6] 山本他：ロボット情報家電コンセプトモデル "ApriAlpha"の開発一機能概要および移動制御について，第 21 回日本ロボット学会学術講演会予稿集，1E27 (2003).
- [7] 大坊：しぐさのコミュニケーション，サイエンス社 (1998)
- [8] 子安：心の理論，岩波書店 (2000).
- [9] 神田他：人間と相互作用する自律型ロボット Robovie の評価，日本ロボット学会誌，Vol.20, No.3, pp.1-9 (2002).
- [10] UPnP(Universal Plug and Play):
UPnP Forum ,<http://www.upnp.org/>