

頸髄損傷者のためのマルチモーダルインタフェース ～使いやすさに着目した家電製品の制御方法～

神奈川県総合リハビリテーションセンター 伊藤英一

1. はじめに

重度身体障害者の日常生活を支援する道具のひとつに環境制御装置（以下、ECS）がある。利用者の身体機能に応じたスイッチやセンサなどを操作することにより、呼び鈴や電動ベッド、テレビ、ラジオ、電話、エアコンなど日常生活に必要な家庭電化製品（以下、家電）の集中制御をおこなうものである。高位頸髄損傷による四肢まひ者（以下、頸損者）により利用されることが多く、呼吸スイッチや音声認識を用いたものが主流となっている。また、操作対象である家電の種類や高機能化により制御指令が増加していることから、操作方法に複雑な階層構造を取るシステムが増えている。最近、マルチモーダルインタフェースが解りやすい操作方法として研究されてきた。そこで、頸損者の家電操作においてマルチモーダルインタフェースを応用し、階層構造によらない ECS を試作した。これは、音声と頭の向きにより、解りやすい操作方法を提供するものであり、その概要と評価実験について報告する。

2. 背景と目的

ECS に対する公的助成制度はほとんど無いため、ECS の導入事例は多くないが、導入支援や操作方法に関する要望は多い¹⁾。

音声認識の福祉応用への要望は高く、音声認識を利用した ECS がいくつか製品化されている。しかし、現状の音声認識のみによる機器操作に関しては操作感の不足など問題の指摘も多い²⁾。さらに家電自体が多様化しており、現状の ECS にある機能ではニーズに合わないことも見うけられてきた³⁾。

高機能化した家電を、ECS により効率よく選択操作するためには階層構造を利用しなければならない⁴⁾。しかし、安易に階層構造などの複雑な操作方法を導入すると誤操作が増え、危険を伴う場面も生じてくると予想される。

そこで、マルチモーダルインタフェース⁵⁾を応用すれば、より解りやすい操作環境が提供できると考えた。例えば、複数の家電を操作する環境には複数のリモコンが存在し、どのリモコンが目的とする機器に対応しているのかで混乱する（図 1a 参

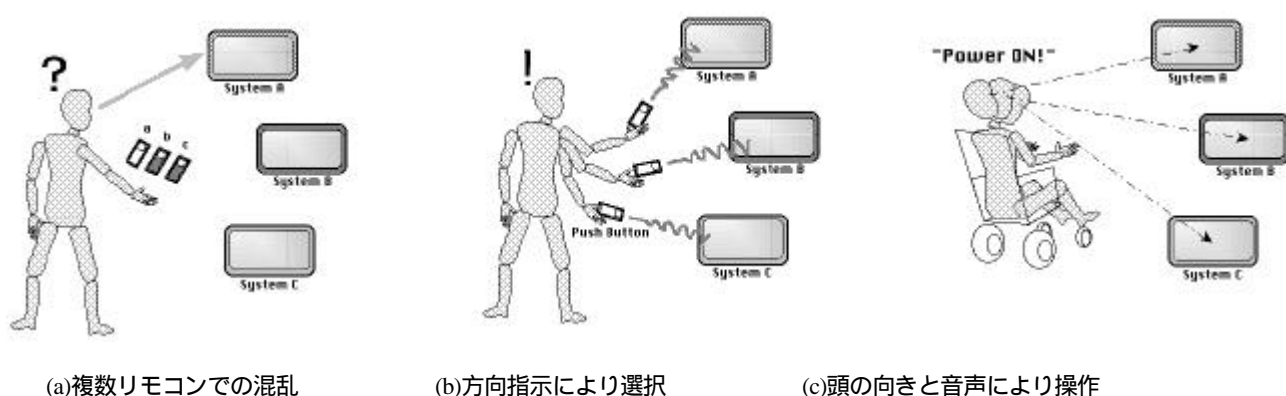


図1 複数の操作対象を操作するリモコンの問題

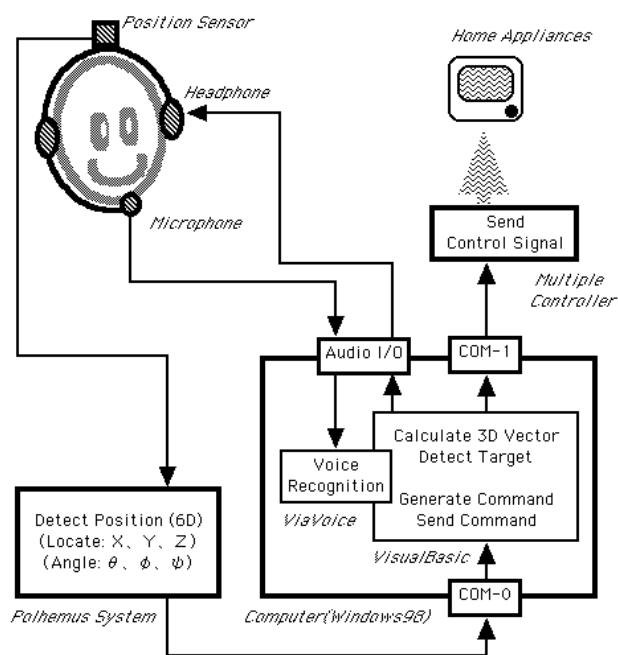


図2 システム構成図

照)。そこに、指差しのような方向指示の動作で操作対象を選択し、共通化されたりモコンで操作できれば、問題は解決できると考えた(図1b参照)。そこで、頭の向きにより操作対象を選択し、音声により制御する操作方法(図1c参照)が頸損者にとって自然で解りやすい手段のひとつとして提案する。

本研究の目的は、新しい ECS の制御手段としてマルチモーダルインタフェースの応用が効果的であると仮定し、その試作と実証実験をおこなうことである。

3. システム概要

研究の目的から考えられるシステムの基本要件は、入力操作としての「頭の向き」と「音声」、および結果としての「家電制御」である。

「頭の向き」を検出する方法として、操作範囲が半径約 1.5m と限定され、かつ接触式(ケーブルが接続)であるため操作範囲は制限されるが、6自由度の検出が可能な磁界による三次元位置計測システム(Polhemus System; ISOTRAK II)を利用する。

「音声」による入力操作としては、口述筆記シ

ステムのため応答性が犠牲となるが入手の容易さから、日本語音声認識(ViaVoicePro)を利用する。

「家電制御」にはエコーネット⁶⁾などのホーム LAN システムが理想的だが、対応する家電の入手が困難、かつ高価であることから学習型マルチリモコン(CROSSAM2+)を利用する。

これらの構成要素を結びつけるための処理ソフトは VisualBasic(VB)で開発し、制御処理用コンピュータ(Windows98SE)で統合処理する。システム構成を図2に示す。

4. 評価実験

4.1 操作環境

ISOTRAK II は磁気を利用するため金属(磁性体)などの影響を受けることから、可能な限り計測範囲内には物を置かないようにした。位置角度センサをヘッドセットに装着したため、頭の動きを制限しないようにケーブルは衣服に固定した。操作対象(以下、ターゲット)は 0.3m 以上離して設置した。音声認識ソフトの多くは話者の特徴検出に必要な前処理(以下、エンロール)が必要であるが、本システムではコマンド入力のみであることからエンロールはしない。

4.2 試用評価実験

被験者は、頸損者(機能レベル C4: 40 歳男性)1名で、電動車いす乗車(図3参照)で操作する。ターゲットは、テレビ、ビデオ、CD ラジカセ、電気スタンド(全てリモコンで操作可)とする。ターゲットへの制御指令としての音声コマンドは事前に利用者へ提示し記憶してもらう。音声コマン



図3 被験者の実験環境

ドとターゲットの組み合わせによる具体的な制御指令を表1に示す。

被験者は、事前に操作訓練を実施した後、(1)音声のみによる操作、(2)頭の向きと音声による操作をおこなった。操作状況はビデオに収録し、音声コマンドや操作内容の経時変化を比較検討した。

表1 ターゲット別制御指令の対応

	テレビ	ビデオ	CD ラジカセ	スタンド
電源	電源ボタン	電源ボタン	電源ボタン	点灯
(数字)	チャンネル		選曲	
音量アップ	音量増加		音量増加	
音量ダウン	音量減少		音量減少	
プラス	音量増加		音量増加	
マイナス	音量減少		音量減少	
点ける	電源ボタン	電源ボタン	電源ボタン	点灯
消す				消灯
プレイ		PLAY ボタン	PLAY ボタン	
ストップ		STOP ボタン	STOP ボタン	
巻戻し		巻戻し	巻戻し	
早送り		早送り	早送り	

5. 結果

(1)音声のみによる操作、(2)頭の向きと音声による操作のそれぞれにおける状況の一例として、正しく「テレビ」の「チャンネル1」を選択した状況における音声コマンドと制御指令のタイミングを図式化したものを図4に示す。実行完了までに費やした時間は、音声のみによる操作では7秒、頭の向きと音声による操作では4.5秒であった。

音声コマンドは12種類と比較的少なく設定したが、錯覚による未設定の音声コマンドを入力することがあった。未設定音声コマンドは入力しても動作しない(未処理)ため、正しい音声コマンドを再入力したことで、誤動作とはならなかった。入力された未設定音声コマンドは「オン」「オフ」「停止」「止まる」であった。

頭の向きによるターゲットの選択は照準が合った場合(ロックオン状態)にサウンド出力でその状態をフィードバックするため、間違えることは無かった。しかし、電動車いすの背部を倒したことにより頭部の方向ベクトルが上昇し、ある背角度以上ではターゲットが視野に入っているものの頭の向きでのターゲット選択ができないという状況が生じた。

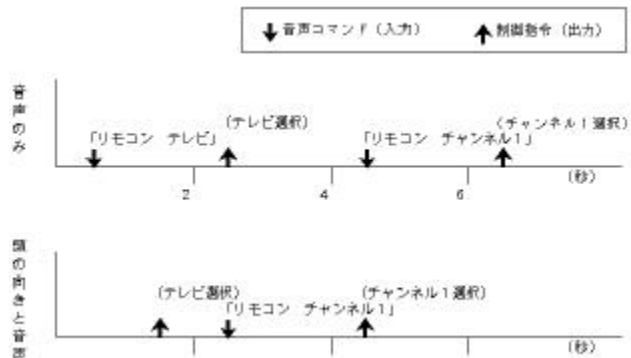


図4 音声コマンドと制御指令のタイミング

6. 考察

音声のみによる操作環境では、ターゲットを選択する指示に加え、そのターゲットを操作するための指示との2つの音声コマンドを用意する方法と、ターゲット毎の操作に対して多数のユニークな音声コマンドを用意する方法がある。

2つの音声コマンドを組み合わせる方法では、ターゲットを特定する音声コマンドと、機器操作のための共通化された音声コマンドを用意すれば良いが、複数の音声コマンドを発声する必要があり、結果で示したように操作時間が延長する。

また、ターゲット毎の操作に対してユニークな音声コマンドを用意する方法では、音声コマンドが『ターゲットの数 x 制御指令』の数だけ必要となり、全てを記憶するか、リストを身近な場所に掲示しなくては利用できない。しかし、ターゲットが増えるほど音声コマンドは増大し、操作は一層困難になると予想される。

頭の向きと音声による操作では、ターゲットを選択するための操作として頭の向きという非言語的なコマンドを利用している。そのため、ターゲットの操作に必要な言語的なコマンドとしては共通化された音声コマンドのみであり、記憶すべき音声コマンドの量も比較的少なく、また1つの操作に必要な音声コマンドは1つであることから時間的にも有効であることが、この実験結果で確認された。

記憶違い、もしくは錯覚による未設定音声コマンドを発声した状況を検討すると、「オン」と「オフ」は反対語であるが、音響的には類似している。もし、このような音声コマンドを設定してしまった場合、誤認識すると動作が反対となる可能性が生じる。このため、音声コマンドとして設定するものと未設定のまま処理しないものとを区別する必要があるのではないかと考えられる。

また、頸損者は起立性低血圧などへの配慮から電動車いすや電動ベッドの背もたれを定期的に倒すことが多い。背角度が倒れている状態では、頭の向きによってはターゲットを示す事が困難となり、操作不能な状況が生じることが確認された。このような状況を想定するならば、背角度などの姿勢に応じてヘッドセットのベクトルに上下方向のオフセット角度を計算することが必要となる。

以上から、本システムは音声のみによる従来の方法よりも有効な手段であり、若干の改善点はあるものの、ターゲットを選択するための言語的なコマンドをなくすことによりコマンド数を減少できたことから、理解しやすい操作環境であると考えられる。

7. 今後の課題

実際の利用のためには、位置角度センサとして利用した ISOTRAK II は計測半径が 1.5m と狭く、また有線式であることに加え、一般的なインタフェースとしては高価であることから、ターゲットの選択方法として赤外線式など別の手段を検討することが急務と考えている。

8. まとめ

ECS として利用するための、「頭の向き」と「音声」によるマルチモーダルインタフェースを用いた、頸損者にとって解りやすい操作システムを試作した。この方法は音声コマンドの共通化によりコマンド数を絞り込む事と、ターゲット選択の指示を非言語化した事により、覚えるべき操作を減らすことができ、また階層構造によらない直接的な操作で時間短縮を図った。また、制御した

いターゲットに頭（顔）を向けることにより、ターゲットの選択が可能となるため、ターゲットの種類が増えたとしても、音声指令の数には影響しない。

頸損者による試用評価実験から、この操作方法的有効性が明らかになった。このように、「音声指令（言語的）」と「頭の向き（非言語的）」のマルチモーダルインタフェースは、四肢まひなど障害のある利用者にとって有効であり、特に今後普及するであろう情報家電のコントローラとして応用されていくことが望まれる。

謝辞

日頃から様々なアドバイスを頂き、また被験者として実験にも立ち会って頂いた金子寿氏に感謝致します。また、システム開発にご協力頂きましたテクノツール（株）梅垣正宏氏に感謝致します。

本研究は財団法人電気通信普及財団平成 11 年度研究助成によるものです。

参考文献

- 1) 横田恒一 他：環境制御装置に関するアンケート調査，第 12 回リハ工学カンファレンス講演論文集，pp.463-466 (1997)
- 2) 太田将主 他：障害者のための音声インタフェースの一考察，第 15 回リハ工学カンファレンス講演論文集，pp.327-330 (2000)
- 3) 山岸秀之 他：頸髄損傷による四肢まひ者の生活と環境制御装置の課題，第 15 回リハ工学カンファレンス講演論文集，pp.605-608 (2000)
- 4) 吉田直樹：選択メニュー学習機能付き環境制御装置の提案と操作効率シミュレーション，第 12 回リハ工学カンファレンス講演論文集，pp.453-458 (1997)
- 5) Cohen, Philip. : Natural Language Techniques for Multimodal Interaction, 電子情報通信学会論文誌 D-II, Vol.J77-D-II, No.8, pp.1403-1416 (1994)
- 6) エコーネットコンソーシアム：http://www.echonet.gr.jp/