

平成 22 年 6 月 10 日現在

研究種目：若手研究（B）
 研究期間：2008～2009
 課題番号：20700471
 研究課題名（和文） 発声を伴った頭部ジェスチャーの認識手法と電動車いすなどの制御インタフェースの開発
 研究課題名（英文） Acoustic-based recognition of the head gesture accompanying voice and its application to a powered wheelchair control
 研究代表者
 佐宗 晃（SASOU AKIRA）
 独立行政法人産業技術総合研究所・情報技術研究部門・研究員
 研究者番号：50318169

研究成果の概要（和文）：手が不自由でかつ明瞭な発声が困難な障害者を対象として、発声時の頭部方向（位置）または発声を伴った頭部ジェスチャーで、電動車いすや環境制御装置など様々な機器の制御を可能にするインタフェースを開発した。これは、センサー類を身に付ける必要のない非拘束・非接触なインタフェースで、前方に置かれた複数のマイクで収録した音響信号から操作者の発声位置を推定することで頭部の位置や動きを検出する。

研究成果の概要（英文）：We have developed a novel interface that enables hand-disabled and speech-handicapped people to control some systems like powered wheelchairs and environmental control systems by means of acoustic-based recognition of the head orientation or the head gesture accompanying voice. The developed interface does not require the handicapped person to wear sensors. Instead, several microphones placed in front of the user acquire the user's utterances. From the acoustical signals, the system can localize the position or trace the movement of the user's head.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008年度	1,800,000	540,000	2,340,000
2009年度	1,600,000	480,000	2,080,000
年度			
年度			
年度			
総計	3,400,000	1,020,000	4,420,000

研究分野：総合領域

科研費の分科・細目：人間医工学、リハビリテーション・福祉工学

キーワード：自立支援技術、マイクロフォンアレイ、音源定位、電動車いす、環境制御

1. 研究開始当初の背景

近年、重度障害者の自立移動の実現に向けて様々な技術開発が行われている。例えば、ヘッドセットマイクを用いた音声認識、磁気セ

ンサーを用いた頭部ジェスチャー認識、そして筋電センサーを用いた筋電信号の認識など、多様なセンサーを用いた電動車いす操作インタフェースが提案されている。しかし、重度

障害者の中にはヘッドセットマイクやセンサー類の装着が困難な場合があり、また、使用中、センサーの位置がずれた場合、自分自身でその位置を修正するのが困難な場合も少なくない。より利便性の高いインタフェースを実現するためには、センサー類を装着する必要のない非拘束・非接触インタフェースを開発する必要がある。このための手段の1つとして、ステレオカメラなどを用いた顔画像処理による頭部ジェスチャー認識がある。しかし、このような手段では、不随意運動のある障害者などの場合、意図しない動作で車いすが誤動作するなどの問題がある。非拘束・非接触インタフェースを実現するもう1つの手段として、本研究課題の提案者は、電動車いすにマイクアレイを実装した音声認識電動車いすの開発を行っている。この音声認識インタフェースではマイクアレイを用いることで、センサー類を身に付けることなく、周囲雑音による誤動作を抑えることが可能となる。しかし、明瞭な音声コマンドの発声が困難な脳性マヒなどの重度障害者では、高い音声認識精度が得られないという問題がある。

2. 研究の目的

この問題を回避するため、本研究課題では、音声認識は行わず、マイクアレイ処理で得られた操作者の発声位置情報に着目したインタフェースを提案する。操作者が発声している時に推定される音源位置は、おおよそ口の位置に一致する。従って、操作者が右を向いて、音声に限らず口笛や舌打ちなど何らかの音を発生したとき、音源位置も右側に推定されるため、操作者が右側を向いていることが検出される。また、何らかの音を発生しながら頭部を、例えば左から右へ動かせば、推定される音源位置も左側から右側へ移動するため、その頭部の動きを検出できる。このようにして得た頭部の向きや動きを認識する

ことで、音声認識を用いない車いす操作インタフェースの構築が可能になる。本研究課題の目的は、操作者が頭部を特定の方向に向けて、もしくは頭部を動かしながら音声に限らず口笛や舌打ちなどの何らかの音を発し、これに対してマイクアレイを用いた音源定位により、操作者の口の位置や移動軌跡を推定することで、頭部の向きまたは音の発生を伴った頭部の動きを検出・認識する非拘束・非接触車いす操作インタフェースを研究開発することである。提案手法は、音声認識を行わないので、発声内容は任意であり、明瞭な音声コマンドを発声することが困難な障害者でも利用が可能となる。



図1 電動車いす搭載型マイクアレイ

3. 研究の方法

(1) 操作者が音を発生しながら頭部を自由に動かすときに、その音源である口の2次元または3次元的位置をより正確に定位できる電動車いす搭載型およびデスクトップ型マイクアレイを開発し、その精度を評価する。

電動車いす搭載型マイクアレイによる音源の2次元定位精度を評価する。図1に評価に用いた電動車いすとマイクアレイを示す。図中、肘掛先端にある黒の四角いスポンジの対角線上にある緑色の基板に、それぞれ4個の

シリコンマイクが実装されている。操作者が車いすに座って、任意の方向を向きながら発声した時の発声位置を2次元的に定位する。それと同時に、操作者は口元に磁気センサーを装着し、その磁気センサーで特定された位置を正解データとして、精度の評価を行う。車いすは屋内だけでなく屋外の様々な環境中を走行するため、音源定位は多様な周囲雑音の干渉に対して頑健に行える必要がある。屋内外合わせて13種類の環境中を、実際に車いすを走行しながら収録した環境騒音を用いて、耐雑音性能を評価する。雑音を混合していない音響信号から定位された位置を基準として、混合雑音のレベルが増加するに従って定位位置の誤差がどのように変化するかを評価する。デスクトップ型マイクアレイを開発し、音源の3次元定位精度を評価する。操作者が発声しながら頭部を自由に動かすときに、その音源である口の3次元的な位置をより正確に定位できるマイクアレイを開発する。このマイクアレイを用いたインターフェースが実用的であるためには、周囲に雑音源がある場合でも、正しく操作者の発声位置を定位できる必要がある。周囲に雑音がない状態で定位した操作者の発声位置を基準にして、雑音のレベルを上げるに従って定位位置がどのように変化するかで耐雑音性能を評価する。

(2) 推定した音源の移動軌跡を高精度に認識できるアルゴリズムを構築する。隠れマルコフモデル(HMM)を用いた手法を開発する。発声を伴った頭部ジェスチャーの発声位置をマイクアレイにより定位する。X-Y座標時系列の学習データを用意し、頭部ジェスチャー毎にHMMを学習する。認識では、観測された座標時系列に対して最も尤度の高いHMMを選択する。高次局所相互相関(HLCC)を用いた手法を開発する。発声位置のX-Y座標時系列からHLCCを求め、それらを

要素とする特徴ベクトルを算出する。HLCC特徴量に対して主成分分析を行い、頭部ジェスチャー毎に特徴量が分布する部分空間の次元数と基底ベクトルを学習する。認識では、観測された座標時系列から求めたHLCC特徴量と、各ジェスチャーの部分空間との距離を算出し、その距離が最小になる部分空間を選択する。これら認識手法の評価として、5名の被験者が発声をしながら、直線的な頭部ジェスチャー9種類と0から9までの数字を描く頭部ジェスチャー10種類の認識実験を行う。4名分のデータで学習し、残りの1名分のデータで認識するオープンテストの評価を行う。

(3) 提案インターフェースを実装した電動車いすや環境制御装置を試作し、障害者に実際に使用してもらうことで評価および改良を行う。発声を伴った頭部ジェスチャー認識による電動車いすの操作インターフェースを開発する。電動車いすの「前進」「右折」「左折」「停止」の操作を、発声時の静的な頭部方向または動的な頭部ジェスチャーで行う。発声を伴った頭部ジェスチャー認識による環境制御システムを開発する。パソコン画面上のカーソル、またはテレビ、ラジオ、ライトなどの家電を、発声を伴った頭部ジェスチャーで制御する。

4. 研究成果

(1) 操作者が発声しながら頭部を自由に動かすときに、その音源である口の2次元または3次元的位置をより正確に推定できる電動車いす搭載型およびデスクトップ型マイクアレイを開発した。電動車いす搭載型マイクアレイによる音源の2次元定位誤差は、左右方向が平均誤差 -0.02cm ($=1.4\text{cm}$)、前後方向が平均誤差 0.16cm ($=2.2\text{cm}$)であった。音源定位は一辺 1cm のグリッド上で行った。図2に、雑音環境下での左右方向と前後

方向の定位誤差を示す。

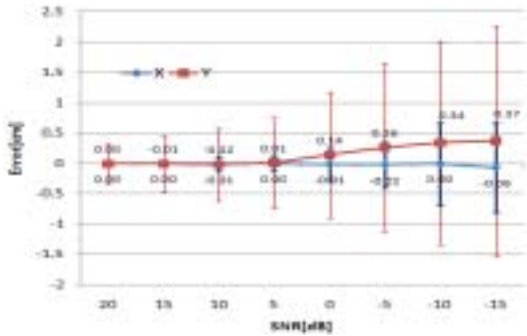


図 2 雑音環境下定位の誤差

音源の 3 次元定位が可能なデスクトップ型マイクアレイを開発した。図 3 に示すように、4 つのシリコンマイクを実装したマイクアレイユニット 3 枚を互いに直交するように配置した構成のマイクアレイを開発した。

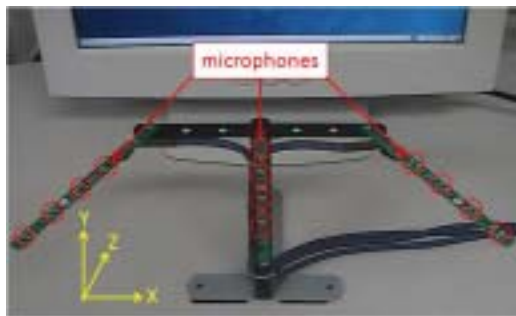


図 3 開発した 3 軸マイクアレイ

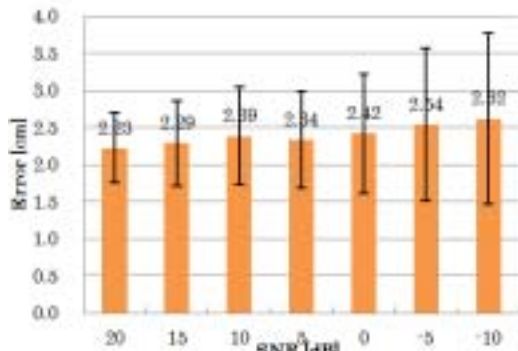


図 4 雑音環境下での音源定位誤差

図 4 は、定位位置が基準位置と異なった場合の誤差の平均値を示す。混合した雑音のレベルによらず平均誤差は 2cm 程度であった。この実験では、音源定位を一辺 2cm のグリッド上で行ったことから、雑音の混合による定位誤差は 1 グリッド程度であったと言える。

(2) 推定した音源の移動軌跡を高精度に認

識できるアルゴリズムを構築した。HMM の状態数を 10 から 30、そして各状態の混合正規分布数を 1 から 4 として、19 種類のジェスチャー認識を行った結果を図 5 に示す。

HLCC の最高次数を 1 から 2、そして相関幅を 2 から 11 として、19 種類のジェスチャー認識を行った結果を図 6 に示す。HMM に比べ HLCC は簡単な処理で実現されるが、最高認識率は 90.25% で HMM の 89.48% を上回った。

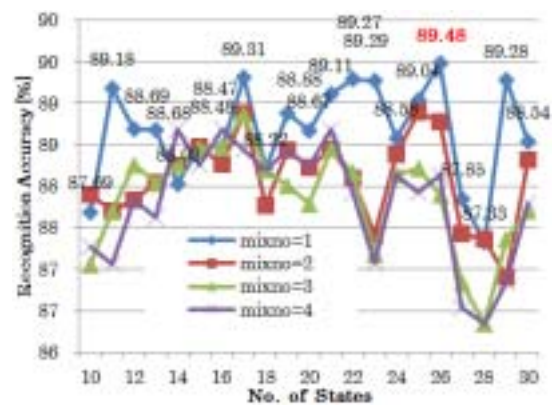


図 5 HMM のジェスチャー認識結果

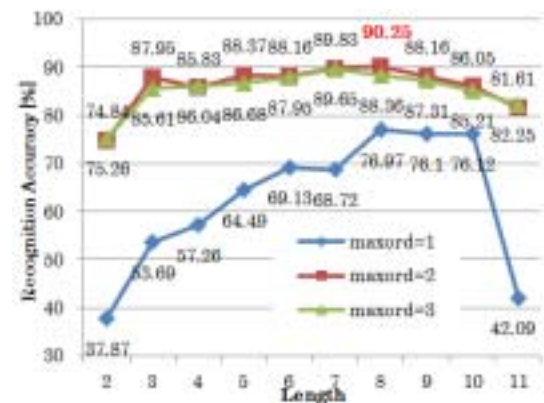


図 6 HLCC のジェスチャー認識結果

(3) 提案インターフェースを実装した電動車いすや環境制御装置を試作し、障害者に実際に使用してもらうことで評価および改良を行った。発声時の静的な頭部方向で操作する電動車いすを図 7 に示す。車いす背後のスピーカーから妨害音が流されている中、安定に操作可能であることを確認した。また発声を伴って頭部を左から右、または右から左へ動か

すジェスチャーをそれぞれ「右折」と「左折」、前傾を「前進」、そして左右に首振りを「停止」とした、頭部ジェスチャーによるインタフェースを開発した。制止状態での発声も含めた5種類の頭部ジェスチャーを5名の被験者を対象にオープンテストで認識した結果、99.67%の認識率が得られた。



図 7 頭部方向で操作する電動車いす

発声を伴った頭部ジェスチャー認識による環境制御装置を図8に示す。図中左下にある橙色のいすに座った操作者の前に、3軸マイクアレイ(図3)が設置されている。マイクアレイを載せている台は取り外し可能で、障害者が日常使用している車いすに設置することも容易である。前方に置かれているディスプレイの左半分の画面にテレビが表示され、右半分に環境制御装置のGUIが表示される。GUIにはテレビ、ラジオ、ライト、それぞれの制御パネルが表示され(図9)、円を描くジェスチャーによりパネルセットが切替る。各パネルには電源のオン・オフや音量操作などの機能が割振られている。パネルセット内の移動は、上下、左右の直線的なジェスチャーで行い、画面に向かって前傾するジェスチャーによりパネルを選択する。この環境制御装置を、手に障害のある障害者に実際に使用してもらったところ、頭部ジェスチャーの認識率は85.7%であった。

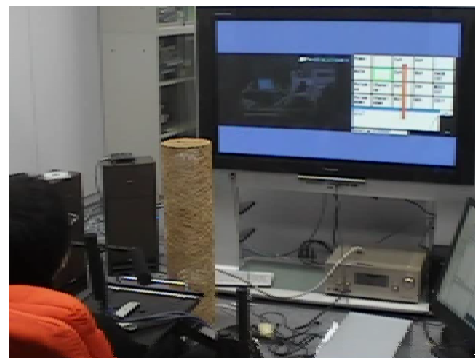


図 8 環境制御装置

Power		1ch	5ch	
mute		2ch	6ch	BACK SET
Volume up	Chanel up	3ch	7ch	NEXT SET
Volume down	Chanel down	4ch	8ch	EXIT

図 9 環境制御装置の GUI

図10に、パソコンのカーソル操作へ応用したシステムを示す。上下、左右方向のジェスチャーによりカーソルを移動する。画面上の小さなアイコンのポインティングを容易に行えるように、前後方向のジェスチャーにより画面上の局所領域のズーム表示機能を実装した。操作者が画面に近づくとサブウィンドウが開き、そこにカーソル周辺を拡大した画像が表示される。カーソルを操作して小さなアイコンをクリックするタスクに要する時間を測ることで、ズーム表示機能の有効性を評価したところ、ズーム表示機能なしでは平均65秒要したのに対し、ズーム表示機能ありでは平均27秒へと短縮された。



図 10 パソコンのカーソル操作

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計1件)

A.Sasou, H.Kojima, "Noise robust speech recognition applied to voice-driven wheelchair", EURASIP Journal on Advances in Signal Processing, 査読あり, Vol.2009, ID 512314, 2009, 19 pages.

[学会発表](計7件)

佐宗晃, 橋本泰治, "高次局所相互相関を用いた発声を伴う頭部ジェスチャ認識", 日本音響学会 2010 年春季研究発表会, 2010/3, 電通大.

橋本泰治, 佐宗晃, "発声を伴った頭部ジェスチャ認識を用いた家電操作インタフェース", 日本音響学会 2010 年春季研究発表会, 2010/3, 電通大.

Y.Hashimoto, A.Sasou, "Development of a 3D Pointing Voice Interface Using a Three-Axis Microphone Array", IEEE intl. sym. on Robot and Human Interactive Communication, 2009/9, Toyama Japan.

A.Sasou, "Acoustic Head Orientation Estimation Applied to Powered Wheelchair Control", Second intl. conf. on Robot Communication and Coordinaton, 2009/3, Odense Denmark.

A.Sasou, "Head orientation estimation integrated speech recognition for the smart chair", Second intl. sym. on Universal Communication, 2008/12, Osaka Japan.

佐宗晃, "発声時の頭部方向により制御可能な電動車いす", 第7回情報科学技

術フォーラム, 2008/9, 慶応大.

橋本泰治, 佐宗晃, 他, "三軸マイクロフォンアレイを用いた三次元ポインティング音声インタフェースの開発", 音声・福祉情報工学研究会, 2008/5, 神戸大.

[産業財産権]

出願状況(計2件)

名称: インターフェイス装置とインターフェイス方法

発明者: 佐宗晃

権利者: 独立行政法人産業技術総合研究所

種類: 特許権

番号: 特願 2008-132542

出願年月日: 平成 20 年 5 月 20 日

国内外の別: 国内

名称: 情報処理装置

発明者: 佐宗晃, 橋本泰治

権利者: 独立行政法人産業技術総合研究所

種類: 特許権

番号: 特願 2008-132552

出願年月日: 平成 20 年 5 月 20 日

国内外の別: 国内

6. 研究組織

(1) 研究代表者

佐宗 晃 (SASOU AKIRA)

独立行政法人産業技術総合研究所・情報技術研究部門・研究員

研究者番号: 5 0 3 1 8 1 6 9

(2) 研究分担者

()

研究者番号:

(3) 連携研究者

()

研究者番号: