

音声認識応用システムの開発

橋場 参生, 中島 康博, 山本 寧

Development of Speech Recognition Applications

Mitsuo HASHIBA, Yasuhiro NAKAJIMA, Yasushi YAMAMOTO

キーワード: 音声認識, 電動車椅子, 自律型走行ロボット, 音声対話インタフェース

1. はじめに

音声は我々にとって最も手軽な情報伝達手段である。このため、機械やコンピュータを音声で操作したいという要望は以前から強く、音声認識技術の実用化に期待が寄せられてきた。

このような背景があったにも関わらず、これまで音声認識技術は、高価な専用装置が必要、認識性能が不十分等の理由から、中小企業の製品開発には応用が進んでいなかった。しかし、近年のハードウェア技術、ソフトウェア技術の急速な進歩によって、不特定話者を対象とする安価な音声認識ソフトウェアや音声認識 LSI が登場し、音声認識技術が実用化段階を迎えつつある。

そこで本研究では、これらの音声認識ソフトウェアや LSI を活用したシステム開発技術の蓄積と道内中小企業への技術移転を目的に、自律型走行ロボットの音声指示システム、音声対話インタフェース、音声操作型電動車椅子の 3 種の応用開発を行った。

さらに、音声操作型電動車椅子については、実用化への開発指針を得ることを目的として、第28回国際福祉機器展等の大規模な展示会に試作機を出展し、意見調査を行った。また、道内の福祉機器メーカーからの支援要請を受け、四肢麻痺の障害を持つ方が使用する特定の電動車椅子に音声操作システムを実装する取り組みを行った。

2. 音声認識ソフトウェアと LSI の動向

2.1 音声認識ソフトウェア

音声認識ソフトウェアは、近年次々と製品化されており、

安価なソフトウェアパッケージとしては、IBM 製の ViaVoice、NEC 製のスマートボイス、スキャンソフト製のドラゴンスピーチ等がある。何れの製品も不特定話者の連続発声に対応しており、音声ワープロ等への有用性がアピールされている。この他にも、東芝、富士通、日立等の大手メーカーが音声認識ソフトウェアを販売している。また、国内研究プロジェクトの成果品として、フリーの音声認識ソフトウェアも提供されている¹⁾。

2.2 音声認識 LSI

従来の音声認識 LSI は、使用者が認識単語を事前に学習させておく特定話者限定単語方式が主流であったが、最近では、不特定話者を対象とする LSI も発売されている。離散形式で発声された音声を認識対象とするため、「ライト・オン」等のように、決まった語句を発声して指示を行うような用途に向くと考えられる。このような音声認識 LSI は、国内では沖電気、NEC、リコー等、海外ではセンサリ社等が販売している。

3. 各種応用システムの開発

音声認識ソフトウェアや音声認識 LSI を活用した応用開発として、自律型走行ロボットの音声指示システム、ファーストフード店等での自動接客を想定した音声対話インタフェース、音声で走行方向を指示できる音声操作型電動車椅子の 3 種の応用開発に取り組んだ。以下、各々の応用開発について報告する。

3.1 自律型走行ロボットの音声指示システム

当場では、オフィスや福祉施設等において、運搬・清掃・案内等の作業を支援・代行する自律型走行ロボットを開発し

事業名: 一般試験研究

課題名: 音声認識応用システムの開発

ている^{2,3)}。この走行ロボットへの命令伝達手段として、音声による指示システムを開発した。図1は赤外線マイクロホンを用いてロボットに指示を送っている様子、図2はロボットのシステム構成である。本システムを利用することにより、走行ロボットに、目的地、作業内容、走行動作等の指示を音声で行うことができる。

図2に示すように、本ロボットは駆動輪の制御や超音波センサ等の信号処理を担当するメインモジュールと、画像処理や音声処理を担当するサブモジュールで構成されており、音声認識機能はサブモジュール内に実装されている。図1に示すように、赤外線マイクロホンに命令を入力すると、ロボットに搭載された赤外線レシーバが音声を受信し、音声認識LSIによって認識処理が行われる。サブモジュールを統括するノート型PCは、認識結果を基に必要な判断を行い、駆動輪の制御が必要な場合には、共有メモリを介してメインモジュールに情報を伝達する仕組みになっている。



図1 自律型走行ロボットの音声指示システム

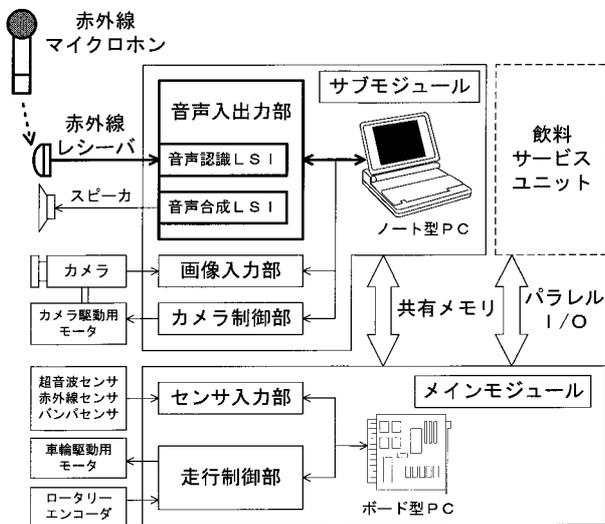


図2 ロボットのシステム構成

3.2 音声対話インタフェース

連続発声を可能にする音声認識ソフトウェアの登場により、対話形式によるコンピュータへの情報入力にも期待が寄せられている。特に、チケット販売やスケジュール予約等、対話内容があらかじめ決まっているような場合には、文法的な制約を加えることによって、認識性能を向上させることが可能である。そこで、このような応用例として、音声認識ソフトウェアと音声合成ソフトウェアを活用し、ファーストフード店での商品注文等を想定した音声対話インタフェースを開発した。その外観を図3に、構成を図4に示す。



図3 音声対話インタフェース

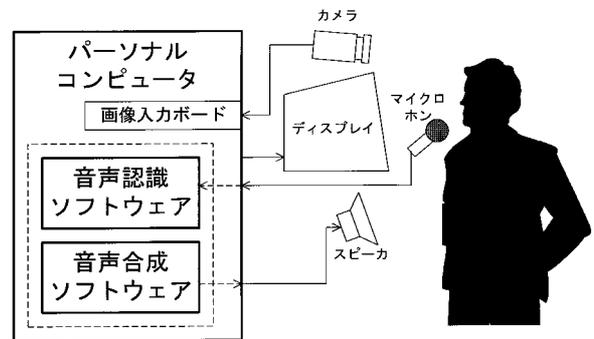


図4 音声対話インタフェースの構成

本インタフェースは、図3の画面が表示されたディスプレイ、マイクホン、カメラで構成されており、カメラの前に利用者が立つと「イラッシャイマセ!。ご注文ラドウゾ。」という合成音声が発せられる仕組みになっている。これに対し、利用者が「ハンバーガー3個とポテト2個とコーヒーを下さい。」等と注文を行うと、商品の確認と、金額の計算が自動的に行われ、結果がディスプレイに表示されると共に、「お会計、××円ニナリマス。」等と合成音声で出力される。

3.3 音声操作型電動車椅子

従来より福祉分野では音声認識技術への期待が大きく、意思伝達装置や環境制御装置等をはじめとして、様々な提案や研究が行われてきた。本研究では、福祉分野への音声認識技

術の応用例として、「前進」・「右回転」等の音声によって走行方向を指示できる音声操作型電動車椅子を開発した。その外観を図5に、構成を図6に示す。



図5 音声操作型電動車椅子

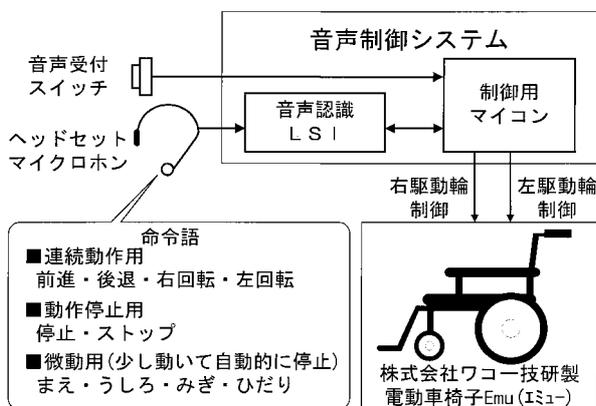


図6 音声操作型電動車椅子の構成

開発に使用した電動車椅子は(株)ワコー技研製のEmu(エミュー)で、これに音声認識LSIを内蔵した音声制御システムを搭載し、音声認識結果に基づいて駆動輪を制御するようにした。搭乗者は、ヘッドセットマイクロホンにより、次のような命令語を用いて走行方向を指示できる。

- 動作開始命令：「前進」「後退」「右回転」「左回転」
- 動作停止命令：「停止」「ストップ」「とまれ」
- 微少動作命令：「まえ」「うしろ」「みぎ」「ひだり」

搭乗者がマイクロホンに向かって動作開始命令を発声すると、指示に従って車椅子が走行あるいは回転する。動作を停止させる場合は、動作停止命令を用いる。微少動作命令は、

停止位置の調整等を想定しており、この命令が入力された場合は、車椅子がわずかに走行あるいは回転した後、自動的に停止する仕組みとした。なお、本車椅子では、雑談等による誤動作を避けるために、音声受付スイッチを用意しており、このスイッチが押されている間のみ、命令を受け付けるようにしている。また、音声受付スイッチをオフにした場合は、動作停止命令の有無に関わらず、車椅子が停止する仕組みにした。図5では、音声受付スイッチを指先で操作しているが、スイッチの形状や取り付け位置を変更することにより、頭部や顎で操作するようにすることも可能である。

4. 音声操作型電動車椅子に対する意見調査

音声操作型電動車椅子に対する障害者や福祉関係者の意見を把握するため、試作機を国内の福祉機器展示会に出展し、デモンストレーションとアンケート調査を行った。本章では、その結果を報告する。

4.1 展示会への出展

音声操作型電動車椅子を出展した展示会は「ヨコハマヒューマンテクノランド」(2001. 6.15-16, 於パシフィコ横浜)と「第28回国際福祉機器展」(2001.10.24-26, 於東京ビッグサイト)である。会場では、多数の来場者に実際に試乗して



図7 ヨコハマヒューマンテクノランドでの様子



図8 第28回国際福祉機器展での様子(中央が試作機)

頂き、95通のアンケートを回収した。出展の様子を図7、図8に示す。

4.2 アンケート結果

アンケート回答者の性別を図9に、年齢構成を図10に示す。回答者の男女比はほぼ同数、年齢は20代から40代が多数を占めるが、80代までの幅広い回答を得ることができた。

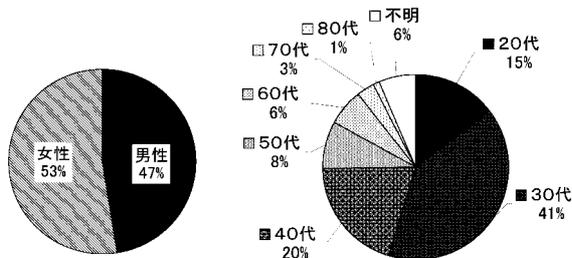


図9 回答者の性別 図10 回答者の年齢構成

まず、「ご覧になった感想はいかがですか?」という設問に対する回答結果を図11に示す。先に記したように、回答者のほとんどは、本車椅子を実際に操作した後、回答を行っている。本設問には、53%が「良い」、40%が「やや良い」と回答し、良好な評価が得られた。

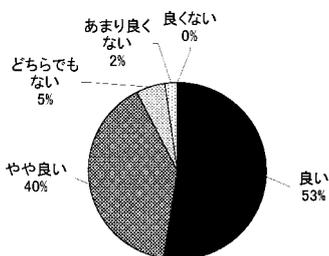


図11 設問「ご覧になった感想はいかがですか?」に対する回答結果

次に、「声で動く車椅子は役に立つと思いますか?」という設問に対する回答結果を図12に示す。本設問には、81%の方が「思う」と回答し、肯定的回答が高い割合を占める結果となった。

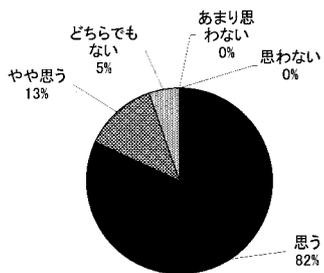


図12 設問「声で動く車椅子は役に立つと思いますか?」に対する回答結果

図13は「車椅子用の音声制御システムが製品化されれば、価格はどの位が良いと思いますか? (車椅子の値段は除いて考えて下さい)」という設問に対する回答結果である。本設問の提示にあたっては、音声制御システムの価格が、電動車椅子本体の価格を上回るのとは現実的でないと考え、30万円未満は5万円刻み、30万円以上については10万円刻みの選択肢を用意し、筆記による回答も可能にした。集計の結果、10万円から15万円未満での製品化を望む回答が最も多く得られた。

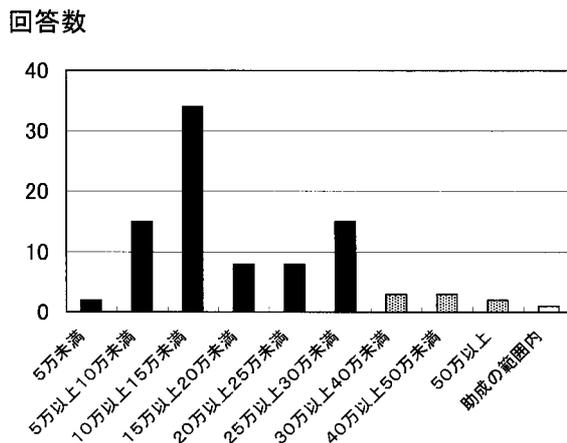


図13 設問「車椅子用の音声制御システムが製品化されれば、価格はどの位が良いと思いますか?」に対する回答結果

続いて、現状の問題点や今後の課題を抽出するために用意した設問「今後、どのような開発が必要だと思いますか?」に対する回答結果を表1に示す。本設問では、表1のa)～e)を複数回答可能な選択肢として用意し、その他の意見がある場合には筆記での回答を依頼した。表1より、約半数の方々が、認識性能の向上、騒音対策、安全性の向上を課題として感じていることがわかった。

表1 設問「今後、どのような開発が必要だと思いますか?」に対する回答結果

今後開発が必要と思う項目	選択率(%)
a) 音声をもっと良く認識できるようにする	56
b) 騒音の影響を受けないようにする	51
c) 緊急停止スイッチをつけるなど、安全性を向上させる	54
d) 前後左右の移動の他にも、もっと色々な操作を行えるようにする	24
e) 音声以外の方法で動かす研究をする	18
その他の回答	
・発声してから動作するまでのタイムラグを短くする。	
・音声受付スイッチを不要にする。	
・誤動作を減らす。	
・車椅子操作中の雑談を可能にする。	など

以上のアンケート結果より、音声操作型電動車椅子は8割以上の方が役立つと考えているが(図12)、現在の試作機を「良い」と評価する方は約半数であり(図11)、実用化に向けては、認識性能、騒音対策、安全性等についての一層の性能向上が望まれていること、また、開発にあたっては15万円程度での製品化を目指す必要があること等が、今後の開発指針として得られた。

5. 音声操作型電動車椅子の実用化

5.1 開発の背景

道内の福祉機器メーカーである(有)たく工房では、障害者の身体機能や体型に合わせたカスタムメイドの車椅子を製作している。同社は、四肢麻痺の方が使用する電動車椅子を製作していたが、使用者の手指機能の低下が著しいことから、ジョイスティックやスイッチに代わる操作方法を模索していた。使用者の発声機能が健全であることに着目した同社は、音声による電動車椅子の操作を実現したいと考え、当场に技術支援を要請してきた。

車椅子の使用者は、札幌在住の18歳の女性で、脊椎脂肪腫という病のために数年前から四肢麻痺を患っている。車椅子使用時は、貧血を避けるために時々背もたれを下げて休息する必要があるとのことから、本機では、前後左右の移動に加えて、リクライニングの操作も音声で行えるようにすることが要求された。

5.2 開発した電動車椅子の概要

開発した電動車椅子の外観を図14に、システム構成を図15に示す。なお、図14はシーティングやカバーリングが行われる前段階の写真である。本車椅子は、背もたれのリクライニング機能を備えた(株)今仙技術研究所製 EMC-30T を使用して開発を行っている。音声制御システムには、試作機と同様の



図14 実用機の外観

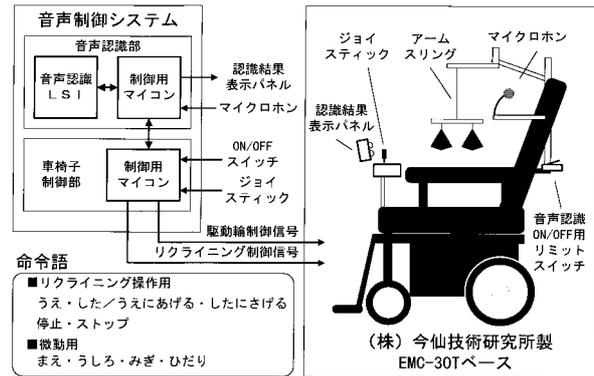


図15 実用機の構成

微小動作命令とリクライニング操作命令を登録し、「うえにあげる」、「したにさげる」等の命令語によって背もたれの上下を操作可能にした。また、ジョイスティックによる通常の操作も可能にし、音声操作と併用できるようにした。

5.3 開発後の経過

本車椅子は、2001年6月に使用者の元に届けられ、実地試験を兼ねて約1年使用されている。この間、音声による操作手順や制御プログラムに改良を重ねることによって、使用者が不具合を感じない程度の性能に至っている。また、現在、使用者はアームスリングの支えがあれば右手でジョイスティックを操作できるようになっており、主に音声によるリクライニング操作機能が活用されている状況である。

6. おわりに

本研究では、音声認識ソフトウェアやLSIを活用したシステム開発技術の蓄積と道内中小企業への技術移転を目的に、各種応用システムの開発を行った。

開発した応用システムのうち、音声操作型電動車椅子に対しては、障害を持つ方々や介護・福祉関係者からの問い合わせも多く、展示会でのデモンストレーションも大変好評であった。また、特定の電動車椅子への実装という形ではあるが、道内企業の製品開発に寄与することができた。福祉分野には電動車椅子の音声操作に限らず、音声認識技術が役立つ事例が多々あると認識しており、今後も積極的に応用開発を進めていきたいと考えている。

一方、音声認識システムの運用にあたっては、マイクロホンの設置方法や環境騒音の状況等により、市販の音声認識ソフトウェア等の活用だけでは十分な性能を得られない場合も多く、実環境における問題を解決するための信号処理手法が必要とされている⁴⁾。従って、今後はDSP等を用いて耐雑音性等に優れた信号処理システムを開発し、既存の音声認識ソフトウェアやLSIと組み合わせしていく取り組みも必要と考えている。

謝 辞

ヨコハマヒューマンテクノランドおよび第28回国際福祉機器展での出展にご協力頂いた株式会社ワコー技研ウエルフェア事業部部长 立川 武氏に感謝致します。

引用文献

- 1) 情報処理学会編集：音声認識システム，オーム社，198PP.，(2001)
- 2) 吉川 毅ほか7名：福祉施設における自律型運搬ロボットの開発（第1報）－全体システムの構成－，北海道立工業試験場報告，No.299，pp.95-104，(2000)
- 3) 吉川 毅ほか7名：福祉施設における自律型運搬ロボットの開発（第2報）－実地評価試験－，北海道立工業試験場報告，No.299，pp.105-110，(2000)
- 4) 松本 弘：音声認識における環境適応技術，電子情報通信学会技術研究報告，99-111，pp.109-114，(1999)