

平成21年 3月31日現在

研究種目：基盤研究（C）
 研究期間：2007～2008
 課題番号：19500476
 研究課題名（和文） 読唇システムを用いた「介護用電動車いす」の走行補助に関する研究
 研究課題名（英文） Research on running assistance of electric wheelchair
 by using lip-reading system
 研究代表者
 小西 亮介（KONISHI RYOSUKE）
 鳥取大学・工学研究科・教授
 研究者番号：00032269

研究成果の概要：

本課題では、手動操作が困難な身障者や高齢者が「口唇の動き」で操作できる「介護用電動車いす」の開発に取り組む。視覚情報を利用して口唇の動きをカメラ画像で認識する読唇の基盤技術、読唇処理のリアルタイム化、センサ情報を利用して障害物などを回避する走行補助に関して、アルゴリズムの開発や電動車いすシステムの実装を行った。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	1,600,000	480,000	2,080,000
2008年度	1,500,000	450,000	1,950,000
年度			
年度			
年度			
総計	3,100,000	930,000	4,030,000

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科学・福祉工学

キーワード：読唇システム

1. 研究開始当初の背景

(1) 高齢化社会において、高齢者が快適で安心して生活を送るためのシステムが提案されている。これらのシステムの普及において、機能が多様化すればするほどそれらの支援システムを簡単に使いこなすことが重要になる。この問題に対して、言葉を利用する方法は、特別な訓練や道具を必要としない利点を持ち、有望なインタフェースと考えられる。

(2) 言葉を利用したインタフェースとして音声認識が提案されているが、周囲雑音により良好な音声認識が行えない問題が未だ解

決されていない。この問題を解決する手段として、視覚情報である口の動きを読み取る読唇がある。

(3) 読唇に関する研究は、国内外で提案されているが、口唇領域の抽出法に主眼したものが多く、実用化まで至っていない例がほとんどである。

(4) 車いす関係の論文として、音声操作型に関する報告は多くあるが、雑音環境等の影響を克服する実用化レベルでの報告はされていない。

2. 研究の目的

本研究の目的は手動操作が困難な身障者や高齢者が「唇の動き」で操作できる次世代型「介護用電動車いす」の開発を目指すことである。この目的を達成するため、下記2課題について取り組む。

(1) 顔を固定した状態で、自然な顔の向きや光源環境にロバストな口唇領域抽出のアルゴリズムおよび読唇の基盤技術を確立する。

(2) 認識結果から車いすを制御するシステムを開発する。認識結果をモータ制御に結び付ける回路を作成し、基礎実験を繰り返し行った後、実際の環境でのテストを行う。

3. 研究の方法

(1) 読唇の基盤技術に関するアルゴリズムの開発

①口唇領域の抽出1：唇輪郭を正確に抽出せず、口内領域のみに着目する。口内領域は唇輪郭と同期して形状が変化する。さらに口内領域の輝度値は概して低く、この性質は光源環境の変化に影響を受けにくい。そこで、輝度値のしきい値処理により口内領域を抽出する。

②口唇領域の抽出2：①により口内領域の抽出を実現できる。しかし、認識精度の向上を図るため、複数の口唇領域抽出に取り組む。これを実現するために Cootes らが提案した Active Appearance Model (AAM) を適用する。

図1は本研究で対象とする顔画像であり、AAMにより抽出された輪郭を描画している。また図2は図1より抽出された五つの口唇領域である。これをもとに特徴量を計測し、認識する。

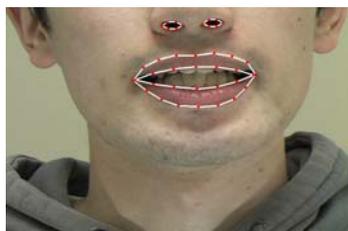


図1 顔画像と輪郭抽出結果



図2 抽出された五つの口唇領域

③五つの口唇領域より各領域の面積やアスペクト比などの13個の形状特徴量を定義する。形状特徴量より、我々の研究グルー

プで提案したトラジェクトリ特徴量 (TF) を求め、DP マッチングによる認識を行う。

④単音、単語など様々な対象に対して認識実験を行う。従来手法と比較することにより提案手法の優位性を評価する。

(2) 読唇処理のリアルタイム処理化

①(1)で開発した読唇アルゴリズムの各処理を統合する。

②事前にビデオカメラ等で撮影した発話シーンでは、発話者はほぼ静止しており光源環境の変化がなく、静的環境であり理想的な顔画像を取得できる。リアルタイム処理では、発話者の顔の傾きやカメラ間の距離の変化により取得できる顔画像が変化する。そのため、これらを考慮して正規化処理を適用する。

顔を動かし、口唇領域の抽出精度や認識精度を評価する。

③リアルタイム処理化に伴い、読唇システムを利用したアプリケーションを開発する。

(3) センサ情報を利用して障害物などを回避する走行補助システムの開発

①読唇システムのメインコンピュータであるノート PC からの信号を受信し、電動車いすを制御する回路を製作する。

②読唇システムとして単語でなく、口の開閉動作と顔の向きを利用して電動車いすを制御するシステムを開発する。

③障害物などを回避するためのセンサとして超音波センサと赤外線センサを搭載する。障害物の位置関係に応じて、減速動作、停止動作、回避動作を行う危険回避システムを開発する。

④③の回避システムを実装した電動車いすとして音声操作型電動車いすを開発する。

4. 研究成果

(1) 読唇の基盤技術に関するアルゴリズムの開発

①日本語5母音に対して認識実験を実施した。被験者は成人健常者10名、各音20回ずつ発話するシーンを撮影した。5母音では時系列データを用いず、最大開口フレーム画像のみを用いた。その結果、94.1%の平均認識率を得た。

②日本語45音に対して認識実験を実施した。被験者は成人健常者5名、各音10回ずつ発話するシーンを撮影した。①の5母音と

異なり時系列データを用いるため TF を利用した。5 母音の分類、10 子音の分類、45 音の分類についてそれぞれ実験したところ、94.7%、30.9%、30.0%の平均認識率を得た。

5 母音への分類は十分な精度を得られているが、子音への分類は低い精度である。この値を評価するため、人間の読唇能力と比較した。40 人の被験者に目視で 45 音の分類を一人 20 回ずつ実施したところ、認識率はそれぞれ 89.3%、28.0%、25.1%であった。この結果より、提案手法の有効性を示した。

③ 日本語の数字 10 単語と車いす制御用 10 単語に対して認識実験を実施した。各実験とも被験者は成人健常者 1 名である。従来手法と比較し、TF と DP マッチングの組み合わせが最も高く、それぞれ 83.5%、99.5%の平均認識率を得た。

④ 従来報告されていなく大規模データ群として、日本語 173 単語を対象とした認識実験を実施した。その結果、92.3%の認識率を得た。

⑤ 従来は特定の言語に対する認識実験しか報告されていない。本研究では日本語、中国語、モンゴル語、ネパール語の 4 ヶ国語、それぞれ 20 単語に対して認識実験を実施した。その結果、どの言語においても TF と DP マッチングの組み合わせが高い認識率を得ることを確認した。

⑥ (1) の研究成果についてまとめると、本研究では従来手法より高い認識精度をもつ認識手法を提案し、その性能を様々な単音と単語に対して検証した。これは国内外の研究者に強いインパクトを与えたと推測する。

文章認識や自然な発話に対する認識、未知単語に対する認識などの課題が残っている。今後はこれらの課題に取り組み、読唇の基盤技術のさらなる確立を図る。

(2) 読唇処理のリアルタイム処理化

① 読唇アルゴリズムを統合し、リアルタイム処理化を図った。開発したシステムの操作画面を図 3 に示す。カメラ画像、領域抽出結果、求められた TF や認識結果を表示する部分から構成されている。

本システムを用いて、顔の動きに対するロバスト性の評価を行った。その結果、ユーザ・カメラ間距離を変えた動きおよび顔の回転動作に対してロバストであることを確認した。

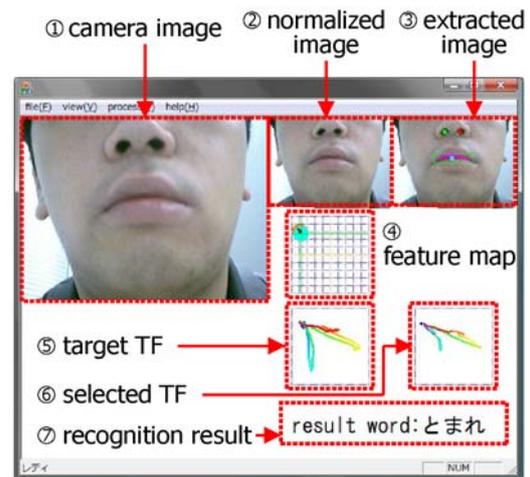


図 3 リアルタイム単語読唇システムの操作画面

② 車いす制御用 14 単語と 47 都道府県の 2 群を認識対象とし、成人健常者 3 名に対して認識実験を行った。その結果、それぞれ 89.5%、82.6%の平均認識率を得た。処理時間は 1 フレームあたり平均 29ms であり、リアルタイム処理化の実現を確認した。発話後の認識時間は、単語数と計算機の性能に依存するが、最も遅い被験者で 14 単語では 1.7 秒、47 単語では 4.6 秒であった。

③ 読唇システムを利用したアプリケーションとして、5 母音の口形認識、口の開閉動作および顔の動きを利用した文字入力システム、意思伝達システムを開発した。

④ (2) の研究成果についてまとめると、従来リアルタイム読唇システムが提案されていたが、その文献では様々な環境における実験が報告されていなかった。本研究では、ユーザや利用環境を変えた実験を行い、実用性のあるシステムの構築を実現した。

本システムは図 1 の顔画像に示すようにユーザの鼻孔と口唇領域の位置がある程度制約されている。そのため、顔の大きな動きが生じると処理が停止する問題が残っている。今後は顔全体が写る状態に対するシステム開発に取り組み、操作の安定性を図る。

(3) センサ情報を利用して障害物などを回避する走行補助システムの開発

① 口の開閉動作と顔の動きを用いて電動車いすを制御するシステムを構築した。構築したシステムの構成を図 4 に示す。

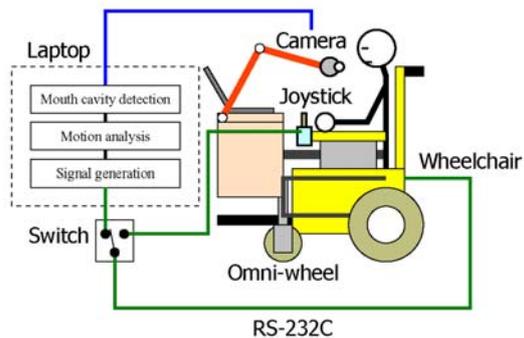


図4 口の開閉動作型電動車いすのシステム構成

大学建物の廊下で成人健常者5人に対して走行実験を行った。約15mの走行コースに対して平均2分26秒の走行時間であった。5人の走行軌跡を図5に示す。二つの障害物を回避して、ゴール地点まで走行できたことが確認できる。

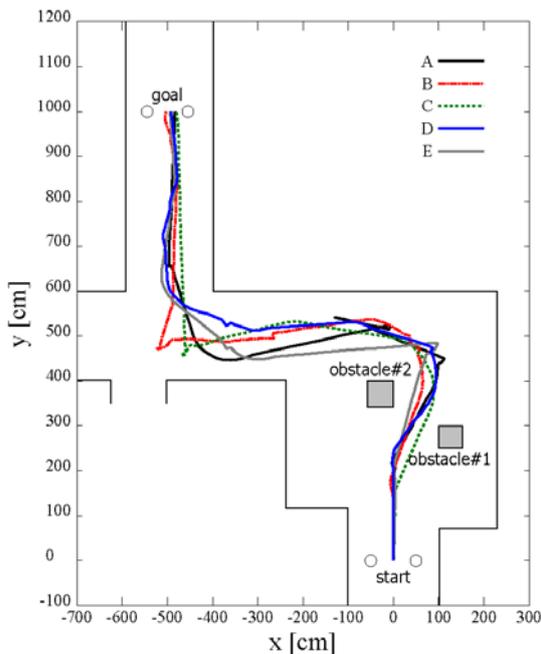


図5 口の開閉動作型電動車いすによる被験者5人の走行軌跡

②①は口の開閉動作を利用しており、インタフェースとしてこれまでにない新規性がある。しかし、廊下の壁や障害物などに対する対処がなされていない。①ではユーザーの操作により障害物を回避しているが、ユーザーに負担を強いる。また障害物に対する対処がないと走行時に危険が伴う。そこで図5に示すような2個の超音波センサと8個の赤外線センサを利用したシステムを構築した。センサのレイアウトを図6に示す。

各センサの計測値より障害物までの距

離を計測し、距離に応じて減速動作、停止動作、回避動作を行う危険回避システムを開発した。またこのシステムの有効性を検証するため、ここでは読唇でなく、音声認識を利用した音声操作型電動車いすを開発した。

さらに、エレベータを利用してフロア間を移動するシステムへの改良を図った。

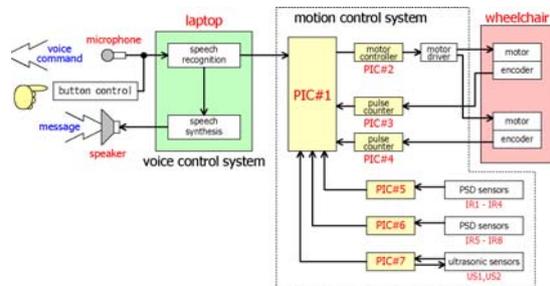


図6 危険回避システムのシステム構成

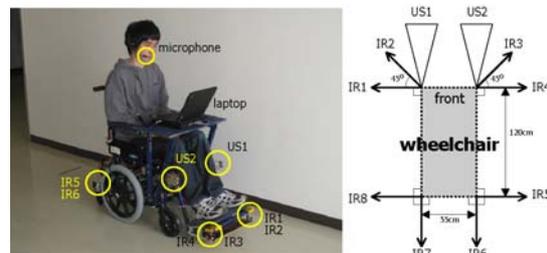


図7 センサの配置図

③②のシステムにおいて成人健常者5人による走行実験を実施した。走行コースは全長約11.9mであり廊下幅が約1.3mの狭いコースと、全長約11.6mであり廊下幅が約2.0倍の広いコース、エレベータを利用して4階から1階へ移動する全長約28mのコースである。危険回避システムを機能させない場合、ジョイスティックによる走行などと比較した。その結果、ジョイスティックに比べると走行時間はかかるものの、衝突することなく走行できることを確認した。ただし、音声操作型電動車いすの走行時間には、ユーザーが命令を入力して車いすが動作するまでの時間が含まれているため、ジョイスティック操作に比べて走行時間が多くなるのは自明である。

④(3)の研究成果についてまとめると、口の開閉動作型を利用した新しいインタフェースをもつ電動車いすの開発、センサ情報を利用して安全な走行が可能な音声操作型電動車いすの開発に取り組み、走行実験により、有効性を示した。

新しいインタフェースを提案したものの操作が難しく利便性に問題が残っている。実用化に取り組むためには操作の利便性は重要な課題である。今後は実用化を目

指した研究にステップアップする。

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 3 件)

① Takeshi Saitoh, Ryosuke Konishi, Japanese 45 Single Sounds Recognition Using Intraoral Shape, IEICE Trans. on Information and Systems, E91-D巻、11号、p2735-2738、2008、査読有

② Takeshi Saitoh, Mitsugu Hisagi, Ryosuke Konishi, Analysis of Features for Efficient Japanese Vowel Recognition, IEICE Trans. on Information and Systems, E90-D巻、11号、p1889-1891、2007、査読有
③ 齊藤剛史, 小西亮介, トラジェクトリ特徴量に基づく単語読唇、電子情報通信学会論文誌、J90-D巻、4号、p1105-1114、2007、査読有

[学会発表] (計 15 件)

① 水口正治、村井彰、西森雅人、齊藤剛史、尾崎知幸、小西亮介、エレベータによるフロア間移動が可能な音声操作型電動車いす、電子情報通信学会福祉情報工学研究会、2009年3月25日、島根大学

② Takeshi Saitoh, Kazutoshi Morishita, Ryosuke Konishi, Analysis of Efficient Lip Reading Method for Various Languages, Proc. of International Conference on Pattern Recognition, 2008年12月8日、米国フロリダ州タンパ市タンパコンベンションセンター

③ 加藤友哉, 齊藤剛史, 小西亮介, トラジェクトリ特徴量に基づく単語認識のリアルタイム処理、電子情報通信学会パターン認識・メディア理解研究会、2008年11月27日、大阪大学

④ 齊藤剛史, 森下和敏, 小西亮介, 発話時における口唇形状の変化による単語認識、計測自動制御学会第13回パターン計測シンポジウム、2008年11月21日、かんぼの宿 修善寺

⑤ 齊藤剛史, 小西亮介, 横顔画像の輪郭形状に基づく読唇、電子情報通信学会パターン認識・メディア理解研究会、2008年9月6日、慶應義塾大学

⑥ 齊藤剛史, 森下和敏, 小西亮介, 複数口唇領域を利用した多言語に有効な単語読唇、電子情報通信学会パターン認識・メディア理解研究会、2008年9月6日、慶應義塾大学

⑦ 齊藤剛史, 久木貢, 森下和敏, 小西亮介, 複数の口唇領域を用いた単語認識、第11回画像の認識・理解シンポジウム、2008年7月29日、軽井沢プリンスホテル

⑧ 水口正治、西森雅人、村井彰、齊藤剛史、尾崎知幸、小西亮介、音声命令による電動車椅子の操作、電子情報通信学会福祉情報工学研究会、2008年5月29日、神戸大学

⑨ 加藤友哉, 齊藤剛史, 小西亮介, リアルタイム口部形状認識を利用した意思伝達システム、電子情報通信学会福祉情報工学研究会、2008年1月25日、島根大学

⑩ 齊藤剛史, 加藤友哉, 小西亮介, 口部パターン形状を利用した文字入力システム、電子情報通信学会パターン認識・メディア理解研究会、2008年1月17日、龍谷大学

⑪ Takeshi Saitoh, Noriyuki Takahashi, Ryosuke Konishi, Oral Motion Controlled Intelligent Wheelchair, SICE Annual Conference 2007, 2007年9月18日、香川大学

⑫ Masato Nishimori, Takeshi Saitoh, Ryosuke Konishi, Voice Controlled Intelligent Wheelchair, SICE Annual Conference 2007, 2007年9月18日、香川大学

⑬ 齊藤剛史, 小西亮介, 唇および口内領域形状に基づくトラジェクトリ特徴量による読唇、第6回情報科学技術フォーラム、2007年9月5日、中京大学

⑭ Takeshi Saitoh, Noriyuki Takahashi, Ryosuke Konishi, Development of an Intelligent Wheelchair with Visual Oral Motion, IEEE International Symposium on Robot and Human Interactive Communication, 2007年8月27日、韓国済州島シンラチェジュホテル

⑮ Takeshi Saitoh, Mitsugu Hisagi, Ryosuke Konishi, Japanese Phone Recognition using Lip Image Information, IAPR Conference on Machine Vision Applications, 2007年5月16日、東京大学

[産業財産権]

○出願状況 (計 1 件)

①名称：文字入力方法

発明者：齊藤剛史、小西亮介

権利者：国立大学法人鳥取大学

特許出願番号：2008-003472

出願年月日：2008年1月10日

国内外の別：国内

6. 研究組織

(1) 研究代表者

小西 亮介 (KONISHI RYOSUKE)

鳥取大学・工学研究科・教授

研究者番号：00032269

(2) 研究分担者

齊藤 剛史 (SAITOH TAKESHI)

鳥取大学・工学研究科・助教

研究者番号：10379654