

令和元年6月20日現在

機関番号：55503

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K21579

研究課題名（和文）ハンズフリー音声認識・視線検出による上肢不自由者への文書作成支援の確立

研究課題名（英文）Establishment of text writing application for upper limb disabled person using hands-free speech recognition and gaze detection

研究代表者

宮崎 亮一（Miyazaki, Ryoichi）

徳山工業高等専門学校・情報電子工学科・助教

研究者番号：40734728

交付決定額（研究期間全体）：（直接経費） 3,100,000円

研究成果の概要（和文）：上肢障害があると、一般のキーボードや携帯電話のデバイスの操作が困難で、メールやレポート等の比較的長い文書を作成するには多大な労力を要する。これらの問題を解決するために本研究では、ハンズフリー音声認識・視線検出システムによる上肢不自由者のための文書作成支援の開発を行った。開発したシステムでは音声認識と視線入力を組み合わせることによって、両手を使用せずに声と視線だけで提示したテキストを完成させることができた。

研究成果の学術的意義や社会的意義

本研究では上肢不自由者に対するテキスト作成支援の基礎を築くことができた。本研究を更に発展させることで、上肢不自由者が健常者と同じように文書やメールの作成を行うことができるようになり、上肢不自由者や介護者のストレスの低減にも繋がることを確信している。また、本研究を発展させることで文書やメールだけでなく両手を使うことなくプログラミングを作成することも期待できる。

研究成果の概要（英文）：Upper limb disabled person is challenging to operate general keyboard and cell phone devices, and it requires a large amount of labor to create long documents such as emails and reports. To solve these problems, in this research, I developed a document writing support application with hands-free speech recognition and gaze detection system. By combining speech recognition and gaze detection, the developed system could complete the presented text with only voice and gaze without using both hands.

研究分野：音響信号処理

キーワード：上肢不自由者 音声認識 視線検出

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

厚生労働省の「生活のしづらさなどに関する調査」(平成 23 年)によると、1 級の上肢不自由者(両上肢の機能を全廃したもの、または両上肢を手関節以上で欠くもの)と 2 級の上肢不自由者(両上肢のすべての指を欠くもの)は 40 万人程度存在する。1 級と 2 級の上肢不自由者は両手の指の機能を欠いているため、健常者と同じようにコンピュータのキーボードの操作を行うことができない。キーボード等を用いて文書やメールを作成できないため、健常者と比べて大きなハンデを背負っていると言える。

上肢不自由者を支援する過去の技術として、視線検出がある。カメラで眼球の動きなどを検出し、画像処理によって視線を検出するものであり、年々検出精度は向上している。しかし、視線検出には文字の入力速度という大きな問題がある。画面上に表示された仮想的なキーボードに視線を合わせて文字を入力する場合、現在の技術では一文字入力するのに 1 秒程度かかる。単語のような文字数の比較的少ない語を入力するような場合においては、視線検出を用いて文字を入力することは有効である。しかし例えば、A4 用紙(約 1200 文字)を埋めるのに要する時間は 20 分程度となる。また、漢字の変換等で文書を作成する時間はさらに増加し、利用者に疲労が溜まるといった問題がある。

一方、近年はフットペダルも開発されている。フットペダルとは、足でペダルを踏み込むことによってキー入力が可能となり、エンターキーや、よく使用するショートカットキーをペダルに割り当てることは有効である。しかし、フットペダルにはキーを数個しか配置されていないため、フットペダルのみで文書を作成することは非現実的である。また、上肢不自由者にとってはフットペダルの持ち運びや設置も大きな問題である。

音声認識システムは近年、Windows や Mac OS 等のオペレーティングシステムやスマートフォンにも導入されるなど、非常に注目を浴びている。また、音声認識精度も静寂な環境であれば、ほぼ 100% 認識することも知られている。しかし音声認識システムの問題点として、「様々な特殊記号の挿入に弱い」、「未知語や固有名詞に対応できない」、「雑音・残響がある環境で使用すると音声認識精度が低下する」などが挙げられる。また、申請者はこれまでに音声認識の性能を向上させるような雑音抑圧の研究を行ってきた。その過程で、現在の音声認識システムでは音声認識性能が 100% に近づいているが、100% になることはないことに着目した。この音声認識の誤認識の問題に対して、一般的にはキーボードを使用して文書の修正や挿入を行うことが前提である。そのため、キーボードが使用できない上肢不自由者においては、文書作成を音声認識システムで代替することは難しいといえる。

2. 研究の目的

本研究は、ハンズフリー音声認識・視線検出システムによる上肢不自由者のための文書作成支援の確立を目指す。上肢障害があると、一般のキーボードや携帯電話のデバイスの操作が困難で、メールやレポート等の比較的長い文書を作成するには多大な労力を要する。現行の音声認識システムは文字の入力速度が速く、認識精度も向上しているものの、雑音や残響の影響により音声認識率は 100% とはならず、誤認識に対しては健常者がキーボード等で修正することが想定されている。しかし、上肢不自由者は上記の誤認識に対する修正ができないといった問題がある。そこで、文書の概形の作成を音声認識で作成し、誤認識を修正するために視線や顔の情報をを用いるシステムを確立し、上肢不自由者が高速に文書を作成できるようになることを目指す。

3. 研究の方法

まず、雑音・残響下で頑健に動作する音声認識システムの構築を行う。本研究の最終目標は、上肢不自由者が実環境下でヘッドセットのような物を着用しないハンズフリー音声認識システムにより文書作成を行うことである。しかし、ハンズフリー音声認識では話者とマイクロホンとの距離が離れるため、目的音声は雑音や残響の影響を大きく受け、音声認識の性能が極端に低下するという問題がある。そのため、一般的には雑音抑圧や残響抑圧等の前処理を施した信号に対して音声認識を行い、音声認識性能の向上を図る。雑音抑圧や残響抑圧に関しては世界中で研究が行われているが、各環境に応じて最高の音声認識性能を引き出すために、内部パラメータを手作業で調整することが多い。音声認識性能の善し悪しは一般的に「SN 比の改善量」、雑音抑圧後の「雑音の品質」・「音声の品質」によって決定されることが知られている。申請者は過去の研究において、雑音抑圧における上記の 3 要素を理論的に導出し、音声認識性能を予測することに成功している。本研究ではさらに背景雑音の種類や残響時間等も音声認識率を予測する要素に加え、音声認識率の予測精度の向上を目指す。さらに、実環境下で背景雑音の種類や残響時間を推定することによって、音声認識性能が最大となる内部パラメータを自動的に最適化するような音声認識システムの構築を行う。

次に、視線検出システムの構築および文書修正システムの開発である。システムはテキスト作成・編集エリアとモード選択に分かれており、音声認識を開始するためのボタン、音声認識システムで誤認識した箇所を修正するための「文字パレット」や、音声認識システムでは入力しにくい記号を入力するための「記号パレット」を表示させることができる。さらに、テキスト作成・編集エリアのカーソルを移動させるための「カーソル移動モード」がある。視線検出システムを構築することによって、モード選択や、文字パレットの文字の選択等、ある程度大

きなボタンを視線によってスムーズに選択する。

4. 研究成果

音声認識システムについては、様々な雑音抑圧手法における音声認識性能を最大限に引き出すパラメータセットの値を求めるために、音声認識実験によって雑音抑圧手法ごとに様々な雑音環境における最適なパラメータセットの値を求め、その値や音声認識率の変化には傾向があるのかということ进行调查した。そして実験結果に対して重回帰分析を用いることで SS と -order MMSE-STSA 法における音声認識性能を最大とするパラメータセットを推定する手法を提案した。比較実験から、提案手法は安定して高い音声認識性能を得られることがわかった。また、雑音環境の推定と、雑音環境に適した音響モデルの構築に取り組んだ。

また、本研究で開発した文書開発システムの外観を図 1 に示す。「パレット」内の「音声認識」ボタンを 2 秒間注視することで Google が提供する Google Cloud Speech API が、「数字認識」ボタンを注視することで大語彙連続音声認識システムを実行する Julius が起動し、「テキストボックス」に音声認識結果が入力される。なお、音声発話を用いた数字の入力を独立して行うために単語辞書を作成し Julius で数字入力を行う機能を実装した。開発したアプリケーションの文字パレットでは「ひらがな」、「カタカナ」及び「濁音文字」の入力が可能で、「入力切替」ボタンを 2 秒間注視することで「ひらがな」、「カタカナ」、「濁音文字」の順でパレットが切り替わる。また、文字を入力する場合にも視線情報を用いる。例えば「あ行」の文字を入力する場合、「あ」のボタンを 2 秒間注視すると、「あ行」の文字(あ、い、う、え、お)が上下左右に展開される。展開された入力したい文字を 2 秒間注視することによって注視した文字が「テキストボックス」に入力される。その後、展開されたパレットは収納され、次の入力待ち状態になる。この手順を繰り返すことで、文字を入力することができる。テキストボックス内の任意の位置へのカーソル移動も視線情報を用いる。「パレット」内にある「カーソル移動」ボタンを 2 秒間注視することでカーソル移動状態になる。その状態で「テキストボックス」内のカーソルを移動させたい任意の位置を 2 秒間注視することでカーソルが挿入される。その後、文字入力状態になり、カーソルを挿入した位置から文字の挿入や削除等ができる。なお、カーソルの微調整は、アプリケーション画面左下に配置している、上下左右の「カーソルキー」を 2 秒間注視することで実現することができる。

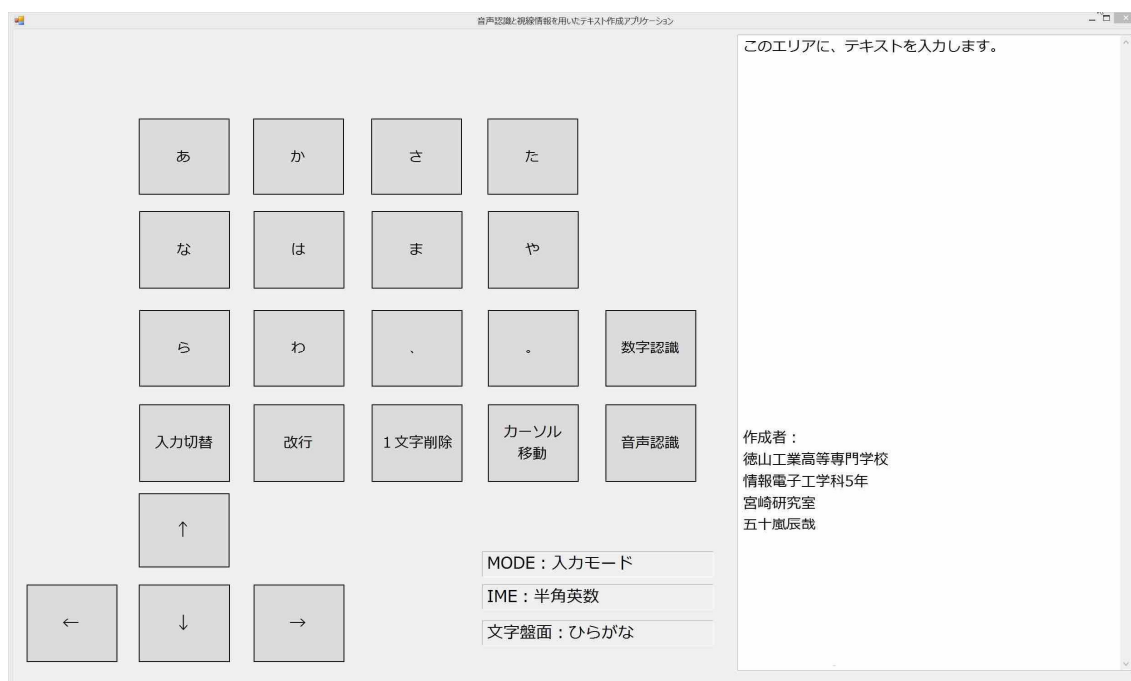


図 1: 文書開発システムの外観

5. 主な発表論文等

〔雑誌論文〕(計 1 件)

宮崎亮一, 五十嵐辰哉, ``音声認識及び視線検出を用いたテキスト作成アプリケーションの開発, ``地域ケアリング, pp.101--103. vol.20, no.11, 2018. (査読有)

〔学会発表〕(計 19 件)

Ryoichi Miyazaki, ``Speech Statistics Estimation for Speech Quality Evaluation in Noisy Environment, ``2018 International Symposium on Intelligent Signal Processing and

Communication Systems, pp.200--204, November 2018.

Masakazu Une, Ryoichi Miyazaki, ``Musical-Noise-Free Speech Enhancement with Low Speech Distortion by Biased Harmonic Regeneration Technique,' ' *International Workshop on Acoustic Signal Enhancement*, pp. 31--35, September 2018.

Shintaro Kubo, Ryoichi Miyazaki, ``Estimation of beta-order MMSE-STSA parameter set for maximizing speech recognition performance with multiple regression analysis,' ' *2018 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP18)*, pp.180--183, March 2018.

Masakazu Une, Ryoichi Miyazaki, ``Evaluation of Sound Quality and Speech Recognition Performance using Harmonic Regeneration for Various Noise Reduction Techniques,' ' *2017 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP17)*, pp.377--380, February 2017.

Shintaro Kubo, Ryoichi Miyazaki, ``Estimation of Spectral Subtraction Parameter-Set for Maximizing Speech Recognition Performance,' ' *5th IEEE Global Conference on Consumer Electronics (GCCE2016)*, pp.567--568, October 2016.

河中昌樹, 宇根昌和, 宮崎亮一, ``入力 SNR を考慮した音声存在確率に基づく単一チャネル雑音スペクトル推定,' ' *日本音響学会講演論文集*, 2-Q-41, pp.427--430, March 2019.

宇根昌和, 宮崎亮一, ``バイアス付き倍音復元技術における内部パラメータと音質の関係性の調査,' ' *応用音響研究会*, pp7--14, December 2018.

河中昌樹, 宇根昌和, 宮崎亮一, ``雑音の過小推定を防ぐ高精度な雑音推定に関する研究,' ' 第 21 回日本音響学会関西支部若手研究者交流研究発表会, December 2018 .

五十嵐辰哉, 宮崎亮一, ``アレー信号処理による音声認識及び視線検出を用いたテキスト作成アプリケーションの開発,' ' 第 94 回福祉情報工学研究会, pp.115--119, March 2018.

垣元宏太, 宮崎亮一, ``一般化ガウス分布仮定に基づく音声カートシス推定法の評価' ' 第 19 回 IEEE 広島支部学生シンポジウム, December 2017.

五十嵐辰哉, 宮崎亮一, ``音声認識と視線情報を用いたテキスト作成アプリケーションの開発,' ' 日本音響学会九州支部第 12 回学生のための研究発表会, pp.23-24, December 2017. (

宇根昌和, 宮崎亮一, ``倍音復元技術に基づくバイアス付き事前 SNR 推定を導入したミュージカルノイズフリー-MMSE-STSA 法の音質改善に関する研究,' ' 音学シンポジウム 2017 (第 115 回 音楽情報科学研究会), no.1, pp.1--6, June 2017.

宮崎亮一, 久保真太郎, ``音声認識性能を最大と雑音抑圧の内部パラメータ推定に関する研究,' ' IEEE CE ソサイエティ西日本合同チャプタ主催 1 月研究会, pp.40--45, January, 2017. (招待講演)

垣元宏太, 宮崎亮一, ``一般化ガウス分布仮定における音声統計量の推定精度に関する評価,' ' 第 19 回日本音響学会関西支部若手研究者交流研究発表会, December 2016.

久保真太郎, 宮崎亮一, ``音声認識性能を最大とする beta-order MMSE-STSA 法のパラメータ推定,' ' 第 19 回日本音響学会関西支部若手研究者交流研究発表会, December 2016.

長尾恭太, 宮崎亮一, ``倍音復元技術を用いたリアルタイム雑音抑圧に関する研究,' ' 第 19 回日本音響学会関西支部若手研究者交流研究発表会, December 2016.

宇根昌和, 宮崎亮一, ``様々な雑音抑圧手法における倍音復元による音声認識性能に関する研究,' ' 第 19 回日本音響学会関西支部若手研究者交流研究発表会, December 2016.

宇根昌和, 宮崎亮一, ``様々な雑音抑圧手法における倍音復元による音声の品質評価に関する研究,' ' 第 18 回 IEEE Hiroshima Student Symposium, pp.402--405, November 2016. (査読有)

久保真太郎, 宮崎亮一, ``音声認識性能を最大とするスペクトル減算法のパラメータ推定,' ' 第 18 回 IEEE Hiroshima Student Symposium, pp.402--405, November 2016. (査読有)

電子情報通信学会ソサイエティ大会, p.58, September 2016.

〔図書〕(計2件)

Hiroshi Saruwatari, Ryoichi Miyazaki, ``Musical-noise-free blind speech extraction based on higher-order statistics analysis (in book chapter),'' *Audio Source Separation (Signals and Communication Technology)*, Springer, pp.333--364, 2018.

宮崎亮一, ``畳み込みってなんですか? (in book chapter),'' *音響学入門ペディア*, コロナ社, pp.8--11, 2016.

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。