

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年 6月17日現在

機関番号：33306

研究種目：基盤研究（C）

研究期間：2010～2012

課題番号：22500519

 研究課題名（和文）聴覚障害学生のための聴講支援システム
 —「質より量」のアプローチによる音声認識—

 研究課題名（英文）Attendance Support System for Hearing Impaired Student
 -- Voice recognition by "Quantity rather than Quality" approach --

研究代表者

川邊 弘之（KAWABE HIROYUKI）

金城大学・社会福祉学部・教授

研究者番号：60249167

研究成果の概要（和文）：異なった認識アルゴリズムとパラメータを持つ音声認識エンジン群をワークステーションで同時に並列実行した。その際の負荷や計算時間、要求される計算機資源などを調べた。また、音声認識エンジンに与えるパラメータを調整し、良好な応答性を与え、妥当な計算量に収まるようにした。低い認識率の音声認識エンジンであっても、種々の音声認識エンジンを数多く同時並列実行し、認識結果に対して多数決を行うことで、少々の誤認識は隠蔽され、結果的に高い認識率が得られた。

研究成果の概要（英文）：I simultaneously executed the voice recognition engines which have different recognition parameters with each other on a workstation and I investigated the load, computation time and computer resources demanded. Then, I adjusted parameters of the recognition engine in order to give good responses. Though many different recognition engines give different results with poor accuracy of recognition, the majority rule hides the error of recognition and finally gives pretty accurate results.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	1,200,000	360,000	1,560,000
2011年度	1,200,000	360,000	1,560,000
2012年度	500,000	150,000	650,000
年度			
年度			
総計	2,900,000	870,000	3,770,000

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：並列処理 / 健康 / 福祉工学 / 聴覚障害 / 聴講支援 / 音声認識

1. 研究開始当初の背景

(1) 本研究に関する研究動向

高等教育機関で学ぶ聴覚障害者が増加している。聴覚障害学生が講義を受講するためには、教員が話す内容を誰かに書き写してもらわねばならない。一般的に聴覚障害学生に対する聴講支援活動は、同じ大学の学生ボランティアによって提供されるのがほとんどである。だが、専門的な講義内容の板書と講師

の話とを同時に筆記することは非常に困難である。また、学生ボランティアが訓練を受け、それなりの技量を身につけたとしても、学生ゆえに4年間で卒業してしまい、技量の蓄積が進まないこともその一因である。それゆえ、通常、複数のノートテイク熟練者が授業に参加し、交互に休息を取りつつ教員の話す内容を要約・筆記する。また、聴覚障害者の受講をサポートするためにインターネッ

トを活用した事例も報告されている。例えば、専門の復唱者による音声入力のインターネットシステムや教室内の画像と音声とを遠隔熟練者へ送り、熟練者は講師の語ったことを書き写して教室内の学生へ送り返し、授業を支援するシステムである。リアルタイムの情報抽出手法を用いた研究としては、電子会議室の発言内容分析するシステムや記述問題の自動評価を目指した教育支援情報システム等が報告されている。だが、これらはノートテイク熟練者や高速ネットワークを求めているため、訓練のための時間や経費が発生し、大学において導入するには困難が伴う。

音声入力システムが普及し始めているが、特定話者向けに調整された場合ですら、その認識率は80%程度であり、ノートテイク熟練者の正確さに及ばない。そのため、音声認識システムを聴講支援活動に本格導入するのは時期尚早と考えられている。

(2) これまでの研究成果

近年、聴覚障害者のために「聴覚障害学生のための聴講システムの研究」に着手した。技量のない学生ボランティアであっても、多人数を投入し、ITで支援することで、聴講支援活動を成立させ得る。熟練者の質を学生ボランティアの量(数)とITとで補うのである。我々はこの考えを「質より量」のアプローチと呼んでいる。このようなシステムならば、大学でも運用できる。

まず、本研究における核となる「多数の低信頼性文字列からの文章の復元問題」に取り組んだ。英語文を対象とした場合、文章の再現性と処理の即時性は良好であり、その成果は国内外諸学会で発表されている。その手法を発展させ日本語文に適用した結果も、国外学会で発表した。また、初心者が講義内容を入力するのに適したデバイスについても検討した。さらに、「質より量」のアプローチの数学モデルを構築し、コンピュータシミュレーションにより、本アプローチの有効性を検証している。

(3) 着想にいたった経緯

我々のシステムの最大の特徴は、記録ボランティアによる入力の未熟さを記録ボランティアの数とITで補うところにある。そこに問題が存在する。①記録ボランティアとして活躍する多人数の初心者学生を確保する必要がある。②キーボードから入力される文章はできるだけ正確なほうが望ましい。

これらの問題を解決するために音声認識に対して、「質より量」のアプローチを適用する。すなわち、キーボード入力する記録ボランティア達を、多数のプロセッサコアで並列実行される種々に調整された音声認識エンジン群で置き換えるのである。音声入力シ

ステムの認識率は、初心者学生ボランティアのキーボード入力における正確さより断然優るが、決して高いわけではない。だが、低い認識率の音声認識エンジンであっても、種々の音声認識エンジンを数多く同時並列実行し、認識結果に対して多数決を行うことで、少々の誤認識は隠蔽され、結果的に高い認識率が得られる。

並列実行はマイクロプロセッサにおける現在の趨勢を反映している。最近のパーソナルコンピュータは2並列ではあるが並列コンピュータとなっている。また、8から16個のプロセッサコアを備えたワークステーションも廉価に市販されている。この状況を考慮すると、音声認識エンジンのアルゴリズムを工夫して認識率を向上させること以外に、多数のプロセッサコアで異なった特徴を持った音声認識エンジンを同時並列実行するアプローチも有望である。(この場合、多数の音声認識結果から最終的な音声認識結果を多数決で抽出することになる。)さらに、我々が行った並列実行のコンピュータシミュレーションでは、10並列程度で95%を超える認識率が得られている。これは、現在のワークステーションの能力で処理できる領域である。

2. 研究の目的

本研究を含む進行中の研究プロジェクトの目的は、聴覚障害学生が大学の講義を不自由なく受講できるシステムの構築である。この研究プロジェクトでは、多数の入力ボランティアのキーボード入力によりノートテイクを実現していた。「質(正確さ)より量(人数)」の概念にもとづいたノートテイクシステムである。そこで問題になったのは、多数の入力ボランティアを確保することと、キーボード入力の正確さであった。そこで、入力ボランティアを確保する問題を解決するため、キーボード入力を講師による音声入力に置き換える。また、音声認識率は約80%と高くはないが、初心者によるキーボード入力より優る。そこで、本研究では、音声認識に「質より量」のアプローチを適用した。すなわち、多数の音声認識エンジンを同時並列実行することにより、音声認識率の向上を目指した。

以上の目的を達成するため、本研究では、以下の3つの目標を設定した。

(1) 音声認識エンジン群の同時並列実行

音声認識エンジン群を1台のワークステーション同時に並列実行する。同時実行数の変化に対する負荷や計算時間、要求される計算機資源などの変化を調べる。また、音声認識エンジンに与えるパラメータを調整し、良好な応答性を与え、妥当な計算量に収まるようにする。

(2) 種々の特徴を持った音声認識エンジンの定義

音声認識システムは認識エンジンと認識エンジンの調節パラメータとから構成される。同一の音声を変えたパラメータで調整された認識エンジンに与えると、異なる認識結果を得る。本研究では、音声認識エンジンに与える調整パラメータの有効範囲を調べ、パラメータを変えることで種々の特徴を持った音声認識エンジンを定義する。目標は10から20程度のパラメータ群を定めることにある。

(3) 多数の低信頼性日本語文字列からの文章の復元

英語文と異なり、日本語での単語分割の困難さが問題を複雑にしている。本研究では、音声認識システム群から出力された複数の日本語文字列データに対して、頻度解析や係り受け解析、特徴分析などを行い、多数決原理を用いて原文復元を試みる。また、復元結果が向上するよう、音声認識エンジンに与えるパラメータを調整する。

3. 研究の方法

本研究では、音声認識エンジンとして **Julius** を用いた。まず、認識アルゴリズムとパラメータが共通な **Julius** システムの複数同時実行から始め、コンピュータシステムへの負荷や計算量、メモリ使用量、認識率などを調査した。この結果により、計算負荷を軽減できるよう認識アルゴリズムとパラメータの調整を行った。これと並行して、その認識アルゴリズムとパラメータを変更することで認識特性の異なる複数の **Julius** システムを作り出した。次に、異なる認識アルゴリズムとパラメータを持つ **Julius** システムの同時実行を行った。ここでも、コンピュータシステムへの負荷や計算量、メモリ使用量、認識率などを調査した。また、認識アルゴリズムとパラメータの調整も行った。そして、**Julius** システム群の出力する日本語テキストに対して、多数決原理を用いて原文復元を試みた。認識アルゴリズムとパラメータの調整、日本語文字列データ解析手法の改良により、認識率を向上させた。

(1) 音声認識エンジン群の同時並列実行

音声ファイルのデータを認識させるのならば、複数のプロセッサを備えた現代のワークステーションで複数の **Julius** を同時に実行することは問題ない。だが、マイク入力された音声を同時実行されている複数の **Julius** に渡すには工夫が必要である。すなわち、リアルタイムで入力されるデータを同時に起動している複数のプロセスに渡すことが求められている。複数の **Julius** の起動で実現した。

また、同時実行数の変化に対する負荷や計算時間、要求されるメモリ量、認識率などの変化を調べた。音声認識エンジンに与えるパラメータを調整し、良好な応答性を与え、妥当な計算量に収まるようにした。

(2) 種々の特徴を持った音声認識エンジンの定義

音声認識システムは認識エンジンと認識エンジンの調節パラメータとから構成される。同一の音声を変えたパラメータで調整された認識エンジンに与えると、異なる認識結果を得る。我々は複数の認識結果に対して、単語単位で多数決処理して最終結果を決めることになるので、その単語認識結果において、すべての音声認識エンジンが誤った認識結果を与えるような状況は望ましくない。どの単語であってでもどれかの音声認識エンジンが正しい認識結果を与えなければならない。さらに、多数決原理で最終結果を決めることを考慮すると、複数の音声認識エンジンが正しい認識結果を与える必要がある。そのような調整パラメータ群が必要なのである。したがって、音声認識エンジンに与える調整パラメータの有効範囲を調べ、パラメータを変えることで種々の特徴を持った音声認識エンジンを定義した。10から20程度のパラメータ群を定めた。

(3) 多数の低信頼性日本語文字列からの文章の復元

日本語での単語分割の困難さが問題を複雑にしている。本研究では、音声認識システム群から出力された複数の日本語文字列データに対して、頻度解析や係り受け解析、特徴分析などを行い、多数決原理を用いて原文復元を試みる。また、分析手法を反映して復元結果が向上するよう、音声認識エンジンに与えるパラメータを調整した。

4. 研究成果

(1) 音声認識エンジン群の同時並列実行

認識アルゴリズムとパラメータが共通な複数の **Julius** をノートPCとワークステーションでそれぞれ同時に並列実行する実験を行った。各CPUコアに1つの音声認識エンジンが割り当てられている間は、音声認識のための実行時間の上昇が低く抑えられ、十分に短い実行時間であった。それ以上の規模の並列実行を行うと、1つのCPUコアを2つまたは3つのプログラムが共同で利用することになるので、実行時間の増える割合が上がってしまい、コンピュータの応答が悪くなった。また、**Julius** のメモリ消費量は、音声認識プログラム1つあたり、約120MBであった。メモリに対する負荷はそれほどではない。以上のことから、我々のノートPCシステムを実

現するには、現在のノート PC はメモリ搭載量では十分だが演算能力では非力で、ワークステーションはメモリ搭載量・演算能力ともに十分だと、結論した。

(2) 種々の特徴を持った音声認識エンジンの定義

音声認識エンジンに与える調整パラメータの有効範囲を調べ、パラメータを変えることで種々の特徴を持った音声認識エンジンを定義した。調整したパラメータは、無音区間の長さ、語彙探索の広さと深さ、日本語辞書の規模、などである。

(3) 多数の低信頼性日本語文字列からの文章の復元

英語文と異なり、日本文での単語分割の困難さが問題を複雑にしている。本研究では、複数の音声認識システム群から出力された日本語文字列データに対して、空白を入れてつづ単語毎に整列を行い、多数決原理を用いて原文復元を試みた。結果として、良好な復元率であった。

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 6 件)

- ① H. Kawabe, K. Sugimori, Y. Shimomura and S. Seto, Voice Recognition System for Hearing Impaired Student --- Error Correction by Principle of Majority Rule ---, Proceedings of 13th Asia Pacific Industrial Engineering & Management Systems Conference, APIEMS 2012, 査読有、2012、1272-1276
- ② H. Kawabe, K. Sugimori, Y. Shimomura and S. Seto, Effectiveness of Simultaneous Execution of Voice Recognition Programs for Hearing Impaired, Proceedings of 12th Asia Pacific Industrial Engineering & Management Systems Conference, APIEMS 2011, 査読有、2012、739-743
- ③ H. Kawabe, K. Sugimori, S. Seto and Y. Shimomura, Simultaneous Execution of Voice Recognition Programs for Note Taking, Everyday Technology for Independent and Care, AAATE 2011, 査読有、2011、401-408
- ④ 川邊弘之, 杉森公一, 瀬戸就一, 下村有子, 音声認識における「質より量」のアプローチの性能評価、金城大学紀要、査読無、Vol. 11、2011、61-68
- ⑤ 杉森公一, 川邊弘之, 瀬戸就一, 下村有子, 聴覚障害者のためのノートテイクシ

ステムの性能評価、日本設備管理学会誌、査読有、Vol. 23、2011、15-23

- ⑥ H. Kawabe, K. Sugimori, Y. Shimomura and S. Seto, Mathematical Model and Computer Simulation of Text Reproduction Based upon “Quantity Rather Than Quality” Concept, CIE40 CD-ROM Conference Proceedings, 査読有、2010

[学会発表] (計 8 件)

- ① 川邊弘之、下村有子、瀬戸就一、杉森公一、聴覚障害者のための音声認識システム -多数決原理による認識率の向上-、情報処理学会第 75 回全国大会、平成 25 年 3 月 8 日、東北大学、仙台
- ② H. Kawabe, K. Sugimori, Y. Shimomura and S. Seto, Voice Recognition System for Hearing Impaired Student --- Error Correction by Principle of Majority Rule ---, 13th Asia Pacific Industrial Engineering & Management Systems Conference, APIEMS 2012、平成 24 年 12 月 4 日、Phuket, Thailand
- ③ 川邊弘之、杉森公一、瀬戸就一、下村有子、聴覚障害者のための音声認識システム -多数決原理による認識誤りの訂正-、聴画像電子学会第 1 回視覚・聴覚支援システム研究会、平成 24 年 8 月 1 日、石川県社会福祉会館、金沢
- ④ 川邊弘之、杉森公一、瀬戸就一、下村有子、音声認識エンジンの複数実行の効果、情報処理学会第 74 回全国大会、平成 24 年 3 月 7 日、名古屋工業大学、名古屋
- ⑤ H. Kawabe, K. Sugimori, Y. Shimomura and S. Seto, Effectiveness of Simultaneous Execution of Voice Recognition Programs for Hearing Impaired, 12th Asia Pacific Industrial Engineering & Management Systems Conference, APIEMS2011、平成 23 年 10 月 16 日、Beijing, China
- ⑥ H. Kawabe, K. Sugimori, S. Seto and Y. Shimomura, Simultaneous Execution of Voice Recognition Programs for Note Taking, 11th European AAATE conference、平成 23 年 9 月 2 日、Maastricht, Netherland
- ⑦ 川邊弘之、杉森公一、瀬戸就一、下村有子、「質より量のアプローチ」による音声認識のシミュレーションと計算機の性能評価、情報処理学会第 73 回全国大会、平成 23 年 3 月 2 日、東京工業大学、東京
- ⑧ H. Kawabe, K. Sugimori, Y. Shimomura and S. Seto, Mathematical Model and Computer Simulation of Text

Reproduction Based upon “Quantity
Rather Than Quality” Concept、The
40th International Conferences on
Computers & Industrial Engineering、
平成 22 年 7 月 27 日、淡路島、兵庫県

6. 研究組織

(1) 研究代表者

川邊 弘之 (KAWABE HIROYUKI)

金城大学・社会福祉学部・教授

研究者番号：60249167

(2) 研究分担者

なし

(3) 連携研究者

なし