

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 12 日現在

機関番号：13501

研究種目：基盤研究(B)（一般）

研究期間：2014～2016

課題番号：26282049

研究課題名（和文）音声処理技術と言語処理技術を活用した電子ノート作成支援システムの研究

研究課題名（英文）Study on an Electronic Note-Taking Support System using Speech and Language Processing Technologies

研究代表者

西崎 博光（NISHIZAKI, Hiromitsu）

山梨大学・総合研究部・准教授

研究者番号：40362082

交付決定額（研究期間全体）：（直接経費） 12,600,000円

研究成果の概要（和文）：本研究では講義等で利用できる電子ノート作成支援システムを開発し、これを用いた際の学習効果を実証することを目的としている。具体的な研究項目は、（1）音声・画像を包括的に記録できる電子ノート作成支援システムの開発、（2）授業音声の記録とノートコンテンツを有効利用するための音声処理・言語処理基盤技術の開発、（3）本システムの学習効果の実証実験、の3点である。

研究成果として、最新の音声・言語処理基盤技術を組み入れた電子ノート作成支援システムの開発に成功した。システムの実証評価結果より、従来の紙のノートよりも使いやすく、かつ電子ノートを用いたほうが学習効果が高いことを明らかにした。

研究成果の概要（英文）：This study aims at developing an electronic note-taking support system for classroom lectures and evaluating it in real environments. In particular, we investigate the learning effectiveness for a lecture when a system user uses the system. We mainly focused on three research topics for the achievement of the goal of this study as follows: (1) development of a note-taking support system which can record a lecturer's talk and capture images projected onto a screen, (2) development of fundamental technologies related to speech and language processing for practical making of note contents and reviewing, (3) investigation of the learning effectiveness when the note-taking system is used.

We successively developed the note-taking system by implementing the state-of-the-art speech and language processing technologies. The experimental results showed that the system was easy to use compared to a traditional note-taking method and the system provided us the high learning effectiveness.

研究分野：音声言語情報処理

キーワード：電子ノート作成支援システム 学習支援 学習効果検証 音声ドキュメント検索 音声認識 音声索引
付け

1. 研究開始当初の背景

日本政府が ICT（情報通信技術）教育の導入を推進しており、「2020 年までに、全国の小・中学校で 1 人 1 台のタブレット端末の整備」という政府目標を掲げている。この流れをくみ、大学等の高等教育機関でもタブレット端末を利用した教育は導入されると予想できる。

研究代表者は、以前に、電子ノート作成支援システムの開発の基礎となる音声メモシステムを開発していた（引用文献①）。この研究では、音声認識技術を利用することで音声認識テキストをメモ手段として活用できること、メモを参照する際に録音された音声を活用することが学習に大変有効であることを示した。このシステムを改良し、教育現場（授業のノート作成手段として）で有効であるかどうかについての予備的検討した。この検討において、音声メモシステムを利用した受講生と紙のノートを利用した受講生とで試験を実施したところ、システムを利用した受講生の成績が良かったこと、紙よりも電子ノートの方が利便性は高いと感じた受講生が圧倒的に多かったことから電子ノートの可能性を見出すことができた。実際に発表聴講者から多くの関心を得たが、本当に学習効果があるかどうか疑問が残っていた。そこで、本研究課題にてこの疑問を追究することとなった。

本研究の特徴は、音声認識技術を利用すること（有効性は確認済）、音声の頭出しや索引付け機能など、音声処理・言語処理を応用した総合的な学習支援までを目指している点にある。近年は音声処理・言語処理の技術革新が活発である。紙のノートの利便さと最先端の音声処理技術・言語処理技術との融合を目指すことで、学習効果の高い電子ノートシステムを実現する。

2. 研究の目的

本研究では、授業で利用するための電子ノート作成支援システムを開発し、これによるノート作成補助および電子ノートコンテンツを学習教材に利用したときの学習効果を検証する。図 1 に研究内容の概要を示す。具体的な研究項目は大きく 3 点である。

(1) 音声・画像を包括的に記録できる電子ノート作成支援システムの開発

誰でも簡単に利用できるための使いやすいユーザインタフェースを設計する。スライド投影の授業スタイルは進行が早い。そのため、ノート記録に集中しすぎるあまり、重要な話を聞き逃すことを防ぐ。このため、音声認識を用いた記録支援を開発し、認識結果をタッチすることで簡単に文字入力を行える機能を開発する。入力される文字は、スライド投影や黒板のキャプチャ画像（教室にあらかじめカメラを設置して、検索・頭出し機能、重要語

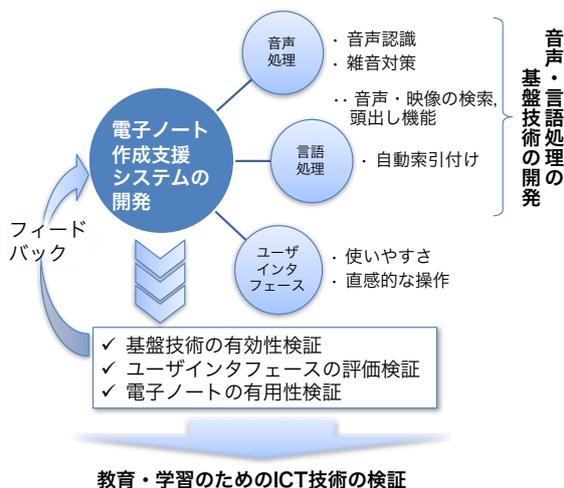


図 1. 研究概要図

句の索引付け機能（書籍巻末の用語索引のイメージ）の実装を行う。

(2) 授業コンテンツを有効利用するための音声処理・言語処理基盤技術の開発

本研究計画では、音声認識技術を利用したノート記録支援が特徴であるため、高い精度で授業音声を書き起こすために音声認識技術の高度化が不可欠である。そのため、授業話し言葉音声の認識率改善のための研究に取り組む。また、前述した検索・頭出し機能、自動索引付け機能において、本研究組織はこれを実現するための基盤技術を有しているが、実環境での利用に対して大きな改善の余地がある。これらの技術をさらに高度化し、基盤技術として確立するための研究開発を行う。

(3) e ラーニング支援として電子ノート作成支援システムの学習効果の実証

(2)で開発した基盤技術を(1)と統合した後、電子ノート作成支援システムの評価を行う。これにより、本システムによる学習効果を客観的・主観的に明らかにする。特に、手書きで書き込むこと、音声認識結果をタッチで書き込むインタフェースはノート作成支援となるのか、自ら作った電子ノート（授業コンテンツ）が e ラーニング教材として有効活用できるのかといったことを明らかにする。

3. 研究の方法

(1) 音声・画像を記録できる電子ノート作成支援システムの設計・開発

研究代表者は、過去に、音声メモシステムを開発済みである（引用文献①）。本研究では、これを拡張する形で開発を行う。電子ノート作成支援システム全体のハードウェア構成を図 2 に示す。発話者（今回の場合、多くは教員・講師）が発話した内容は常にマイクで収録され、音声認識されて字幕としてユーザ端末に送られる。またユーザ（多くの場合は学生）は授業を聞きながら必要に応じてメモを

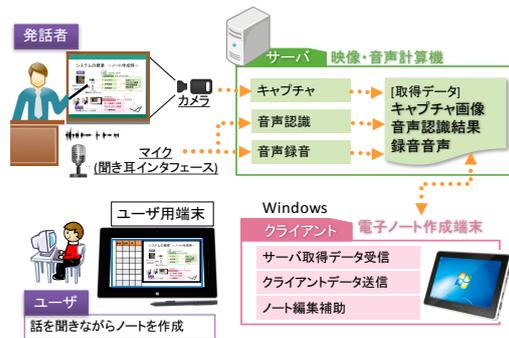


図2. 電子ノート作成支援システムの構成図

とる。このとき、ユーザの任意のタイミングで、カメラで撮影されたスクリーンの画像や黒板の画像を電子ノートに取り込むことができる。

電子ノート作成支援システムの開発では、以下の2つの項目に留意しながらインタフェースの開発を行う。

① 使いやすいユーザインタフェースの開発

音声認識を利用した入力インタフェースによって、認識結果をタッチするだけでメモとして記録が残る。また、手書き入力で任意の文字・図形をメモとして記録できるようにする。これらのメモは、キャプチャされたスクリーン（黒板）画像や事前に配信された授業資料上に自由に書き込めるように設計する。また、ユーザにとってどのような機能があれば便利で使いやすくなるか、といったことを、プロトタイプシステムを開発しながらその都度被験者実験を通して確認していく。

② ノート参照機能（検索・頭出し機能）の開発

録音（録画）された音声・映像を検索するための機能や、必要な箇所を瞬時に頭出し再生できる機能を開発する。音声認識技術を発展させ、ある語句のメモが行われた時刻とそのときに発声されていた音声（映像）との同期をとる技術を開発することで、この機能を実現する。

(2) 音声処理・言語処理基盤技術の開発

① 音声認識の高度化

雑音や残響に頑健な音声認識、歩きながら発声した音声に頑健な音声認識を実現するための音声認識技術を開発する。音声認識を行うためには音声波形データ（代表的なものにWAVE形式）を「特徴量」系列へ変換する。このときに雑音・残響に強い特徴量を模索する。また、深層学習技術を用いた音声認識手法を開発する。

② 映像・音声データの検索・頭出し機能のための基盤技術の高度化

長い授業コンテンツから、任意のフレーズを発声している部分を見つけたり、あるキーワードに関連している内容を検索したりする

には、マルチメディアデータの高度な検索技術が必要である。これには「音声ドキュメント検索」の基礎技術を発展させる。

③ 自動索引付け機能の開発

学習支援のために、作成した電子ノートからの自動的に重要語を判定・抽出し、索引付けを行う技術を開発する。これには研究代表者の持つ音声からのキーワード検出の基礎技術を発展させる。

(3) 電子ノート作成支援システムの評価（被験者実験）

音声処理・言語処理の各基盤技術と電子ノート作成支援システムとを統合し、電子ノート作成支援システムを完成させ、システムの評価を行う。評価実験では、実際に電子ノート作成支援システムの被験者評価を行う。評価項目は、客観的評価（音声認識性能、コンテンツの検索性能、索引付け機能の性能）と、アンケートによるユーザインタフェースの主観的評価（使いやすさ、音声認識性能が使いやすさに与える影響はどうか？等）とする。さらに、電子ノート作成支援システムを用いる効果（ノート作成時およびeラーニングで利用する学習効果）を調査するため、電子ノートを使わなかった被験者と、実際の試験で比較する。従来の学習方法（紙のノート等）を利用した被験者群と比べて、開発した電子ノート作成支援システムを利用した被験者群の学習理解度（ペーパー試験を実施）の有意な改善を目指す。

4. 研究成果

(1) 音声・画像を記録できる電子ノート作成支援システムの設計・開発

図3に本研究で開発した電子ノート作成支援システムのグラフィカルユーザインタフェースを示す。音声認識結果を表示する部分と、スクリーン（黒板）のキャプチャ画像を貼り付ける部分に分かれている。キャプチャはノート使用者の任意のタイミングで貼り付ける

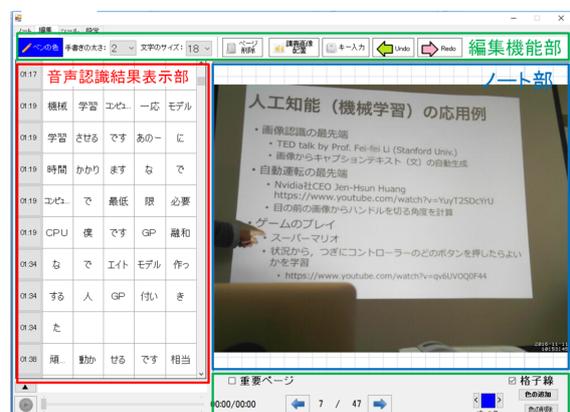


図3. 開発した電子ノート作成支援システムの主要画面

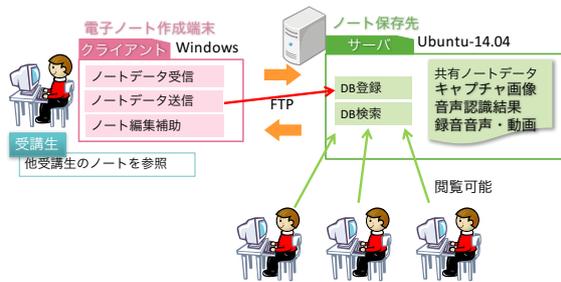


図 4. 電子ノートの共有機能

ことができる。画像の上にはキーボードや手書き文字によってメモを記入できる他、音声認識結果の単語をなぞることで、なぞられた単語をメモとして利用できる。

図 4 は本システムで作成したノートコンテンツを共有する仕組みを示した図である。個人個人で作成したノートコンテンツをサーバ上で共有することで、他人のノートを閲覧できる。また、講師も受講生がメモした内容を参照できるため、講義の参考にもなる。

図 5 は、音声検索・索引付け機能のためのユーザインタフェースである。作成したノートコンテンツには講師の話した音声すべてが保存されている。ノート参照の際に、すべての音声を聞き直すことが困難な場合、指定したキーワードに関する内容について説明してある箇所を特定し、その箇所から聞くことができるようになってきている。例えば、図 5 に示すように、『音声認識システム』というキーワードで検索を行うと、音声の索引付け結果から素早くキーワードが含まれている場所を特定できる。図 5 の●をクリックしても、その部分の頭出しが可能となっている。

(2) 音声処理・言語処理基盤技術の開発 主な成果は以下の通りである。

① 音声認識の高度化

講義音声の高精度な認識のための講師の話者適応技術を開発した。従来は音声分析には三角窓フィルタを用いているが、新しくガウシアン窓・ガンマトーン窓によるフィルタバンクを深層ニューラルネットワーク (DNN) に組込むことで、少数の学習サンプル数でも話者に適用した音響モデルを学習できることを示した。

また、ビデオカメラを利用することから、映像を伴うシーンの音声認識を高精度化するための技術を提案した。DNN を用いた映像と音声それぞれからの特徴抽出結果を統合することにより音声認識の高精度化を実現した。

さらに、深層学習による音声認識誤りの自動訂正技術も開発し、高い音声認識精度を実現した (音素認識率 98%)

② 音声検索・索引付け技術

電子ノート作成支援のために、音声検索語検出を発展させて、確信度の順に検出結果を利用し、漸進的に検索結果を提示する音声内

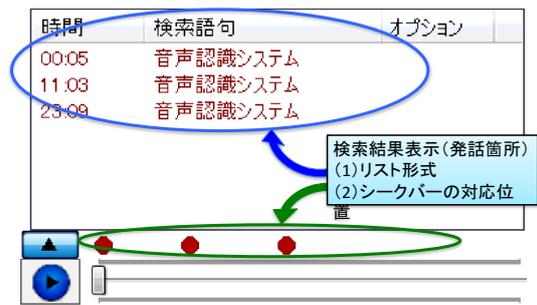


図 5. 音声検索・索引付け機能

容検索手法を開発した。また、ノート利用者の自由発話音声を検索要求とする音声内容検索を検討し、音響特徴を利用した検索手法を開発した。

(3) 電子ノート作成支援システムの被験者評価実験結果

主に 2 つの観点で開発した電子ノート作成支援システムの被験者実験を行った。これらについて述べる。

① 音声認識機能・音声検索機能の調査

「3. 研究の方法」の(2)項で開発した音声認識技術や音声データの検索機能が電子ノート作成支援システムに与える各種の影響を調査した。

まず、音声認識率がどの程度、電子ノート作成支援システムに役に立つのかを調査した。その結果、音声認識率が高いに越したことはないが、音声認識率が 60% あれば利用できるという結果であった。今回開発した音声認識システムを使うと、被験者実験対象のざっくりとした話し言葉の講義に対しては約 60% の認識精度であった。スライド投影のキャプチャ画像を有効活用できるという点から、そこまで講義音声の書き起こし提示に対する高い必要性はないとの意見もあった。なお、音声認識率を意図的に劣化させた場合 (約 30% の認識精度) に対しては、まったく役に立たないという結果となった。

音声認識率が 60% とさほど高くなくても (認識誤りが起きたとしても)、人間は認識誤りが起きていることを認知し、誤っている単語から正しい単語を想起できる能力がある。したがって、少々誤った単語が混在していても内容はある程度理解できており、例えば講師の話の聴き逃したとしても、認識結果を見ればすぐにリカバーできるという点で、認識結果の提示は役立つ。

また、授業で作成した電子ノートの参照時の利便性を評価するために、音声検索・索引付け機能 (図 5) の有用性の調査を行った。これは、電子ノートで作成した授業コンテンツを見ながら授業の試験問題を解くという評価である。このときに、音声検索機能があった場合となかった場合とで、ノート参照の利便性を客観的に評価した。客観データとしては、

試験問題を解き終わるのに要した時間とした。実験の結果、音声検索機能を使った被験者グループの平均解答時間は35分であり、使わなかった被験者グループの平均解答時間は40分であったことから、電子ノート参照時、図5のような音声データを検索する機能は大いに役立つとの結果となった。多くの授業は約90分の長さがあり、効率的に復習するには収録されている講師の説明を素早く探すことが必要であり、これを支援する上で音声検索機能は役に立つことが明らかとなった。

② システムの有用性・学習効果の検証

作成した電子ノート作成支援システムの評価を行った。

実際に開講した4つの講義すべてに参加する大学生・大学院生12名に、被験者として電子ノートシステムを利用いただいた。講義1回目～3回目終了後に小テストを実施し、4回目の講義終了にはシステムの有用性をアンケートで評価してもらった。小テストでは作成したノートの内容を見ても構わないこととした。また、4回目の講義終了後の2ヶ月後に、再度同じテストを実施することで学習の定着の評価を行った。

まず、小テストの結果では、電子ノートシステム利用者の平均点が89点、その他紙のノートをとるなど電子ノートシステムを利用していない受講者の平均点は63点となり、有意に成績が改善していることが分かる。電子ノートシステムによって、講義中の集中度が高まり、記録漏れも減ることが示唆された結果となっている。

2ヶ月後に再度同じ問題のテストを実施したところ、電子ノートを利用していた受講者の平均点は64点で、システムを使わなかった受講者の平均点は45点であった。このことから、電子ノートシステムを使って講義ノートを作成したほうが、授業内容の定着していることが明らかとなった。

さらに、電子ノートシステムについての主観評価結果については、紙のノートと比べ：

- ・紙のノートよりも聞き逃すことが減った
- ・講義により集中できた
- ・見直しがずいぶん楽

となり、紙のノートよりも電子ノートを使うほうが受講生の学習を支援できそうだということが明らかとなった。

(4) 成果のまとめ

最後に本研究課題の成果を簡潔にまとめる。

① 基盤技術として以下の研究開発を行い、有効性を実証した。

- ・雑音に頑健な音声認識技術（雑音除去や映像と音声との併用技術）
- ・話者に依存した音声認識技術
- ・音声認識結果に対する誤りの訂正技術
- ・音声データに対する高精度な検索・索引付け技術

② 電子ノート作成支援システムの有用性を評価し、以下の事柄を明らかにした。

- ・従来の紙のノートよりも使いやすい
- ・音声認識結果を提示することでメモがとりやすくなることはもちろん、聴き逃しにも対応できる
- ・画像にメモを書き込むだけで良いので、授業に集中できる
- ・復習時（見直し時）にすぐに必要な箇所の情報（説明音声など）を閲覧できる
- ・電子ノートを使ったほうが紙のノートをとるよりも学習効果があり、記憶に定着しやすい

<引用文献>

① 西崎博光, 太田晃平, 関口芳廣, “聞き耳メモリー：音声認識結果を取捨選択して利用する電子メモ作成支援アプリケーション”, 電子情報通信学会論文誌, Vol. J96-D, No.10, pp. 2379-2389, 2013

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 4 件)

- ① Naoki Sawada, Hikomitsu Nishizaki, “Re-Ranking Approach of Spoken Term Detection using Conditional Random Fields-based Triphone Detection,” IEICE Trans. on Information & Systems, 査読有, Vol.E99-D, No.10, pp.2518-2527, 2016. DOI:10.1587/transinf.2016SLP0012
- ② Kentaro Domoto, Takehito Utsuro, Naoki Sawada, Hikomitsu Nishizaki, “Spoken Term Detection using SVM-based Classifier Trained with Pre-indexed Keywords,” IEICE Trans. on Information & Systems, 査読有, Vol.E99-D, No.10, pp.2528-2538, 2016. DOI:10.1587/transinf.2016SLP0017
- ③ 西崎博光, 胡桃澤圭佑, 西崎香苗, 池上仁志, “携帯端末で録音された音声メモを用いる診療記録作成支援システム”, 電子情報通信学会論文誌, 査読有, Vol.J99-D, No.3, pp.358-366, 2016. DOI: 10.14923/transinfj.2016JDP7074
- ④ Aditya Arie Nugraha, Kazumasa Yamamoto, Seiichi Nakagawa, “Single-channel dereverberation by feature mapping using cascade neural networks for robust distant speaker identification and speech recognition,” EURASIP Journal on Audio, Speech, and Music Processing, 査読有, Vol.2014, pp.1-31, 2014. DOI: 10.1186/1687-4722-2014-13

[学会発表] (計 4 件)

① 成田陽介, 西崎博光, “電子ノート作成支援システムの利用が及ぼす学習効果の検証”,

- 情報処理学会第 79 回全国大会講演論文集,
Vol.4,2ZE-04,pp.881-882,2017 年 3 月 16
日発表,名古屋大学(愛知県・名古屋市)
- ② 中村卓磨,澤田直輝,西崎博光:“音素遷移
ネットワークを用いたリアルタイムキー
ワードスポッティングの検討”,日本音響
学会 2016 年秋季研究発表会講演論文集,
2-Q-11, pp. 65-68, 2016 年 9 月 15 日発表,
富山大学(富山県・富山市)
- ③ 山田一星,西崎博光:“講義音声認識のため
の言語モデル学習ユーザインタフェース
の設計”,日本音響学会 2016 年春季研究発
表会講演論文集, 3-P-16, pp. 195-196,
2016 年 3 月 11 日発表,横浜桐蔭大学(神
奈川県・川崎市)
- ④ 南條浩輝,高橋徹,西崎博光:“初等教育授
業音声の利活用のためのアーカイブ技術
の基礎的検討”,日本音響学会 2016 年春季
研究発表会講演論文集, 3-P-17, pp. 197-
200, 2016 年 3 月 11 日発表,横浜桐蔭大
学(神奈川県・川崎市)

[図書] (計 2 件)

- ① 日本音響学会(編集),秋葉友良 他,コロナ
社,日本音響学会編 音響キーワードブ
ック, 2016, 494
- ② 寺嶋一彦(監修),中川聖一 他,情報機構,
今後の超高齢化社会に求められる生活支
援ロボット技術, 2015, 609

[その他]

ホームページ等

<http://www.alps-lab.org/kikimimi/>

6. 研究組織

(1) 研究代表者

西崎 博光 (NISHIZAKI, Hiromitsu)

山梨大学・総合研究部・准教授

研究者番号: 4 0 3 6 2 0 8 2

(2) 研究分担者

秋葉 友良 (AKIBA, Tomoyosi)

豊橋技術科学大学・工学研究科・准教授

研究者番号: 0 0 3 5 6 3 4 6

北岡 教英 (KITAOKA, Norihide)

徳島大学・社会産業理工学研究部・教授

研究者番号: 1 0 3 3 3 5 0 1

中川 聖一 (NAKAGAWA, Seiichi)

豊橋技術科学大学・リーディング大学院研
究推進機構・特任教授

研究者番号: 2 0 1 1 5 8 9 3

宇津呂 武仁 (UTSURO, Takehito)

筑波大学・システム情報工学研究科・教授

研究者番号: 9 0 2 6 3 4 3 3