

令和 3 年 6 月 22 日現在

機関番号：21201

研究種目：基盤研究(C)（一般）

研究期間：2018～2020

課題番号：18K11358

研究課題名（和文）方言音声のビデオアーカイブ化と方言音声理解のための情報処理技術の確立

研究課題名（英文）Video Archiving of Dialect Speech and Establishment of Information Processing Technology for Understanding Dialect Speech

研究代表者

伊藤 慶明 (Yoshiaki, Itoh)

岩手県立大学・ソフトウェア情報学部・教授

研究者番号：90325928

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究課題は、遠野地域の語り部の会からの要望もあり、語り部の遠野昔話の語りの収録を行い、13人約14時間の方言音声を含むビデオデータが構築でき、遠野方言・民話保存活動に関する地域貢献を果たした。また、標準語とは言語・音素・イントネーション体系が異なる方言音声に対して検索技術の確立を図り、語彙に依存しない音声検索システムを適用し、高精度・高速・低メモリでの検索方式を研究開発した。さらに、一般の人が遠野方言の民話を聞いても理解できるように語りの理解をサポートできるように、上記技術を適用し、語り部が話している民話中のキーワードに対しリアルタイムの検出を理論上可能であることを確認できた。

研究成果の学術的意義や社会的意義

失われつつある方言音声を保存する点において言語学的にだけでなく、音声言語情報処理学的にも学術的特色を有するとともに、遠野地域の語り部の民話保存活動に貢献しその社会的意義が大きい。方言音声には音声認識技術が利用できなかったが、我々の語彙によらない音声検索方式を応用し、方言音声の検索機能の実現は学術的意義を有する。方言音声の聞き手の内容理解を支援するアプリケーションには社会的意義が大きいと考える。

研究成果の概要（英文）：In response to a request from a group of storytellers in the Tono area, this research project recorded the narration of Tono folktales by the storytellers. Video data were constructed containing approximately 14 hours of dialectal speech by 13 people. The project made a contribution to the community regarding the preservation activities of the Tono dialect and folktales.

We established a search technology for dialectal speech that differs from the standard language in language, phonemes, and intonation system. In addition, we applied the above technology to support the understanding of the narrative so that the general public can understand the folk tales of the Tono dialect, and confirmed that it is theoretically possible to detect keywords in the folk tales spoken by the narrator in real time.

研究分野：音声処理

キーワード：方言音声処理 音声中の検索語検出 音声認識 深層学習

1. 研究開始当初の背景

近年、テレビやインターネットの普及に伴い、地方でも通常の会話で標準語が話されるようになり、方言を話す人および地域が縮退傾向にある。一方、方言を守る・残す機運も高まっており、遠野「語り部」1000人プロジェクトのように方言・民話の保存と語り部を受け継いでいく活動が行われている。遠野市では十数年前に語り部の民話の保存のためビデオを収録したが、現在の語り部の話す民話の保存については費用・人手の面から進んでいないのが現状で、遠野文化研究センターでは早急な収録を望んでいた。

本研究課題の研究代表者（東京出身）と大学4年生の学生（兵庫出身）が、「とおの物語の館」で毎日開催されている語り部の民話を実際に聴取したが、方言が強く民話中の重要なキーワードさえ聞き取るのは難しく、内容が良くは分からず、残念な経験をした。この点についても遠野文化研究センターでは、苦慮しているところであるが、語り部の話している民話に字幕をつけてしまうと、方言が持つ温かさ・面白さが軽視される恐れがあるため好ましくなく、キーワード的なものが表示・解説されるような仕組みが望まれた。

2. 研究の目的

上記のような問題認識のもと本研究課題では、我々のこれまでの研究成果を踏まえ以下の目的のもと、研究課題に取り組む。

① 遠野方言民話のビデオアーカイブ化

現在、遠野の語り部は10年前に収録した数名の語り部の音声しか残っていない。そこで本研究課題では、現在の語り部によるなるべく多くの民話の収録を新たに行い、遠野方言・民話保存活動に関する地域貢献を果たす。

② 方言音声の検索機能の実現

近年深層学習の進展により、音声認識が高度化し実利用もされ始めている。一方、方言については対応できていないのが現状である。数百万人が話す関西弁でさえその音声認識システムは構築されていない。方言には辞書やその言葉の並びをテキスト上で記述されたデータがないため、音声認識システムを構築することは不可能に近く、日本国内、海外でも方言に対する研究が進んでいない状況にある。一方、我々がこれまで研究を進めてきた、語彙に依らない音声検索システムは、音節や音素をベースとして検索を行うため、方言への適用が期待できる。音素体系も異なる方言音声に対して、本研究ではまず、方言音声に対する検索技術の確立を目指す。

③ 検索技術に基づく民話の理解・方言の理解のためのサポート技術の確立

一般の人が遠野方言の民話を聞いても前述の通り理解しきれないのが現状である。折角の語り部の話す、味のある、温かみのある遠野方言の語りを十分伝えきれていないのは、遠野文化研究センターでも危惧するところである。本研究課題では、この語りの理解をサポートする技術を先の検索技術をベースに確立する。これにより、語り部の話を遮ることなくまたその温かみも維持しつつ、聞き手の理解のサポートができると考える。

3. 研究の方法

これら2つの科研費テーマの研究開発実績および、ビデオデータベースの構築実績を踏まえ、以下3つのサブテーマに分けて研究開発を推進する。図1に本研究のイメージを示す。

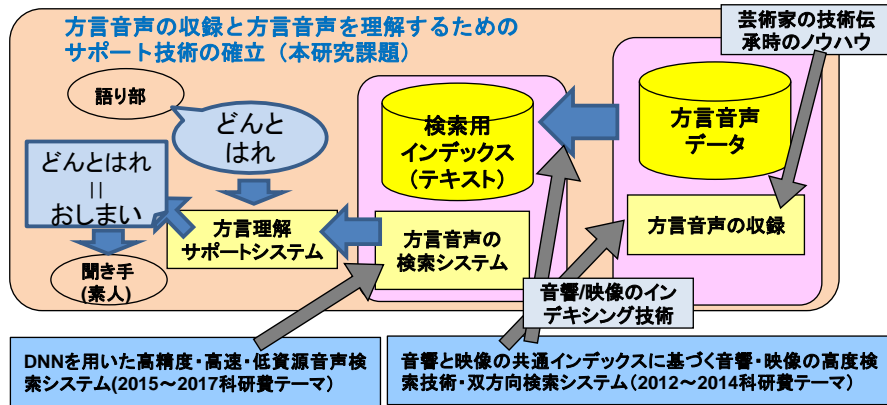


図 1. 本研究課題のイメージ図

- ① 方言音声の収録： 遠野文化研究センターと協力して遠野地方の語り部が話す民話の方言音声を 10 人各 10 話程度ずつ約 20 時間収録し遠野方言・民話保存活動に関する地域貢献を果たす。その際には、芸術家の技術伝承のた図 1 本研究テーマのイメージめのビデオアーカイブ化のノウハウを活用しながら効率的な収録技術の確立を目指す(図右)。
- ② 方言音声の検索システムの研究開発：

方言には辞書やテキストで記述されたデータがなく、単語をベースとした音声認識システムを利用できない。前々回科研費テーマの音響・映像情報のインデキシング技術と前回科研費テーマの音声検索システムを応用し、音声認識システムの音素体系も異なる方言音声に対して、語彙に依存しない音声検索システムの応用を図る。音声認識システムの音響的特徴モデルと方言の音響的特徴とは異なるため、最新の深層学習を用いて方言への新しい適応化技術・比較照合方式の研究開発を行う。まずは、方言音声の情報処理技術の確立の第一歩として、入力が簡便な音声で検索語を入力し、語り部が話している民話中のキーワードに対し検索を可能にするシステムの研究開発を行う。次に、方言音声に対してテキストで検索語を与えた場合にも最新の深層学習技術の一つである CTC (Connectionist Temporal Classification)の技術を応用し、音声認識システムの辞書にない任意の言葉、すなわち方言や未知語に対して検索可能なシステムの研究開発を行う。
- ③ 方言理解サポートシステムの研究開発：

方言で民話が話されている最中に、その民話を標準語化した字幕を表示しても、民話の温かさが失われ遠野文化研究センターでは望ましくないと考えている。一方、上述の方言音声の検索システムでは聞き手が分からない部分やキーワードをその場で判断・理解する助けとはならない。民話を理解する上で重要なキーワードが民話中に出現したときに、その意味を聞き手に提供できれば、聞き手は民話の内容を理解しやすくなる。そこで本研究課題では、上述の検索システムを拡張し、民話の重要キーワードをスポット的に検出し、リアルタイムに聞き手の理解を支援するシステムを研究開発する。

4. 研究成果

本研究テーマは 3 つのサブテーマにより構成されており、研究テーマを各サブテーマの研究成果を以下に各年度ごとに示す。

(1)2018 年度の研究成果は以下の通り。

本研究課題では、2018 年度は主に現在の語り部による民話の収録を新たに行い、遠野方言・民話保存活動に関する地域貢献を果たすことと、標準語とは音素体系が異なる方

言音声に対して検索技術を確立すること、また、一般の人が遠野方言の民話を聞いても理解できるように語りの理解をサポートする技術を確立すること、これらを目的として、3つのサブテーマごと以下の研究開発を実施した。

① 方言音声の収録：遠野文化研究センターおよび遠野昔話語り部の会と協力して、遠野地方の語り部が話す民話のビデオ収録を開始し、現段階で10人約10時間の方言音声を含むビデオデータが構築できた。遠野方言・民話保存活動を進めている遠野地域への地域貢献を果たすことができた。

② 方言音声の検索システムの研究開発：方言には辞書やテキストで記述されたデータがなく、単語をベースとした音声認識システムを利用できない。標準語の音声認識システムの音素体系・単語体系が異なる方言音声に対して、既に研究開発を行っている語彙に依存しない音声検索システムの応用を図った。音声認識システムの音響的特徴モデルと方言の音響的特徴とは異なるため、最新の深層学習モデルを方言音声に適用した。まずは、方言音声の情報処理技術の確立の第一歩として、語り部が話している民話中のキーワードに対し検索・検出を可能にするシステムの研究開発を行った。また、方言音声の検索に、キーワードをテキストあるいは音声で与えた場合の検索方式の検討を行い、検索に適切な音声言語単位を評価した。

その結果、音声で検索キーワードを与える場合よりもテキストで検索キーワードを与える方が高い検索精度が得られること、また、テキストで検索する場合、音声データを処理する言語単位を単語、音節等があるが、これらを複数並列して用いて検索し、結果を統合することでより高い精度が得られることを確認した。

(2) 2019年度の研究成果は以下の通り。

本研究課題では、2019年度は、現在の語り部を対象として遠野昔話の収録を継続して行い、遠野方言・民話保存活動に関する地域貢献を果たすことと、標準語とは音素体系が異なる方言音声に対して検索技術を確立すること、さらに、一般の人が遠野方言の民話を聞いても理解できるように語りの理解をサポートする技術を確立することを研究目的として以下の研究開発を推進した。

① 方言音声のデータベース化：遠野文化研究センターおよび遠野昔話語り部の会と協力して、遠野地方の語り部が話す民話のビデオ収録を完了し、13人約14時間の方言音声を含むビデオの収録が完了した。遠野方言・民話保存活動を進めている遠野地域での利用方法について語り部の会とともに検討を実施した。

② 方言音声の検索システムの研究開発：一般に方言には辞書がなく、方言音声にテキストが付与されたデータがなく、標準語の音声認識システムを利用できない。標準語の単語体系・音素体系と異なる方言音声に対して、我々の開発した語彙に依存しない音声検索システムの適用を行った。本年度は、検索精度を向上させるため、最新のLong short term memory(LSTM)等の深層学習モデルの利用を図った。また、方言音声の検索に、英語などの多言語を用いることにより検索精度の向上を図った。これらの検索技術を応用し、語り部が話している民話中のキーワードに対しリアルタイムに検索・検出を可能にするシステムのフェージビリティスタディを行った。

その結果、LSTM等の最新の深層学習モデルを複数併用することで高い検索精度が得られること、日本語以外の言語の音声認識システムを用いることで検索精度が向上する可能

性があることを確認できた。さらにリアルタイムにキーワード検索を実現するために、Posteriorgram を集約することで高速化できることが確認できた。

(3) 2020 年度の研究成果は以下の通り。

最終年度である 2020 年度は、遠野地域の語り部の会からの要望である語り部の遠野昔話の語りの収録したデータをデータベース化し、遠野方言・民話保存活動に関する地域貢献を果たすこと、および標準語とは言語・音素・イントネーション体系が異なる方言音声に対して検索技術を確認すること、さらに、一般の人が遠野方言の民話を聞いても理解できるように語りの理解をサポートする技術を検討すること、これらの当初目的を研究の目的として、に以下の研究開発を推進した。

- ① 方言音声のデータベース化：遠野文化研究センターおよび遠野昔話語り部の会と協力して、遠野地方の語り部が話す民話のビデオ収録を完了し、13 人約 14 時間の方言音声を含むビデオデータが構築できた。遠野方言・民話保存活動を進めている遠野地域へ無料で提供し、遠野地域での活用が可能となり地域貢献を果たすことができた。現在、方言音声の研究利用方法についての検討を開始した。
- ② 方言音声の検索システムの研究開発：一般に方言には辞書がなく、方言音声にテキストが付与されたデータがなく、標準語の音声認識システムを利用できない。標準語の単語体系・音素体系と異なる方言音声に対して、我々の開発した語彙に依存しない音声検索システムの適用を行った。これまでに、検索精度を向上させるため、最新の Long short term memory(LSTM)や双方向 LSTM, End-to-end モデル等の深層学習モデルの利用を図った。また、方言音声の検索に、英語などの多言語を用いることにより検索精度の向上を実現した。
- ③ 語り部が話している民話中のキーワードに対しリアルタイムに検索・検出を可能にするシステムに向けて、高精度・高速・低メモリでの検索方式の研究開発を推進した。Posteriorgram 照合では高い検索精度が得られるが、速度が遅く、メモリ量が膨大であるため、実用化は困難であることから、これまで開発した、最尤状態系列化方式の検索技術を応用し、深層学習モデルの出力ノード数を少数に限定した深層学習モデルを複数利用することで、高精度・高速・低メモリでの検索が実現できることを確認できた(引用文献 2)。これにより、語り部が話している民話中のキーワードに対しリアルタイムの検索・検出の実行可能性が検証できた。

[1] 伊藤慶明, 岩崎瑛太郎, 金子大祐, 小嶋和徳, 李時旭, “音声中の検索語検出における音声クエリ・音声ドキュメントのフレームレベル最尤系列化照合方式,” 電子情報通信学会論文誌 D, Vol.J100-D, No.8, pp.798-807, (2020/09/01) Online ISSN: 1881-0225, DOI: 10.14923/transinfj.2016JDP7122.

[2] 西野将弘, 伊藤慶明, 小嶋和徳, 李時旭, “異種・複数の深層学習モデルを用いた音声中の検索語検出方式の高精度・低メモリ化” 日本音響学会春季研究発表会, 2-2Q-4, pp.1079-1082 (2021-3).

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 丹治 遥, 小嶋和徳, 李時旭, 南條浩輝, 伊藤慶明	4. 巻 61
2. 論文標題 音声中の検索語検出におけるクエリに関連語を利用したリスコリング方式	5. 発行年 2020年
3. 雑誌名 情報処理学会論文誌	6. 最初と最後の頁 103-112
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 伊藤慶明, 岩崎瑛太郎, 金子大祐, 小嶋和徳, 李時旭	4. 巻 Vol. J100-D, No.8
2. 論文標題 音声中の検索語検出における音声クエリ・音声ドキュメントのフレームレベル最尤系列化照合方式	5. 発行年 2020年
3. 雑誌名 電子情報通信学会論文誌 D	6. 最初と最後の頁 798-807
掲載論文のDOI (デジタルオブジェクト識別子) 10.14923/transinfj.2016JDP7122.	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計17件（うち招待講演 0件 / うち国際学会 3件）

1. 発表者名 Takashi Yokota, Kazunori Kojima, Shi-wook Lee, and Yoshiaki Itoh
2. 発表標題 Reduction of Speech Data Posteriorgrams by Compressing Maximum-likelihood State Sequences in Query by Example
3. 学会等名 Asia-Pacific Signal and Information Processing Association APSIPA (国際学会)
4. 発表年 2020年

1. 発表者名 Haruka Tanji, Kazunori Kojima, Hiroaki Nanjo, Shi-wook Lee, and Yoshiaki Itoh
2. 発表標題 A Rescoring Method Using Web Search and Word Vectors for Spoken Term Detection
3. 学会等名 Asia-Pacific Signal and Information Processing Association APSIPA (国際学会)
4. 発表年 2019年

1. 発表者名 丹治遥, 小嶋和徳, 李時旭, 南條浩輝, 伊藤慶明
2. 発表標題 SQ-STDのフレームレベル状態系列間照合における格助詞を用いた検索精度改善方式
3. 学会等名 日本音響学会春季研究発表会
4. 発表年 2020年

1. 発表者名 横田平志, 小嶋和徳, 李時旭, 伊藤慶明
2. 発表標題 音声中の音声検索語検出のための最尤状態系列化方式の音声データのPosteriorgram圧縮法
3. 学会等名 日本音響学会春季研究発表会
4. 発表年 2020年

1. 発表者名 飯田英仁, 伊藤慶明, 小嶋和徳, 李時旭
2. 発表標題 遠野物語における複数言語の音声認識モデルを用いたキーワード検索精度向上
3. 学会等名 情報処理学会第81回全国大会
4. 発表年 2020年

1. 発表者名 西野将弘, 伊藤慶明, 小嶋和徳, 李時旭,
2. 発表標題 音声中の検索語検出における状態数の異なる複数の深層学習モデルを用いた検索精度の向上
3. 学会等名 情報処理学会第81回全国大会
4. 発表年 2020年

1. 発表者名 横田平志, 伊藤慶明, 小嶋和徳, 李時旭,
2. 発表標題 音声中の検索語検出における音声クエリ最尤系列化方式のメモリ使用量削減方式
3. 学会等名 第2回東北地区音響学研究会
4. 発表年 2019年

1. 発表者名 西野将弘, 伊藤慶明, 小嶋和徳, 李時旭
2. 発表標題 “音声中の検索語検出におけるCTCとBLSTMを用いた検索の低メモリ高精度化
3. 学会等名 第2回東北地区音響学研究会
4. 発表年 2019年

1. 発表者名 Shi-wook Lee, Kazuyo Tanaka and Yoshiaki Itoh
2. 発表標題 Empirical Analysis of Score Fusion Application to Combined Neural Networks for Open Vocabulary Spoken Term Detection
3. 学会等名 INTERSPEECH 2019 (国際学会)
4. 発表年 2018年

1. 発表者名 丹治遥, 小嶋和徳, 李時旭, 南條浩輝, 伊藤慶明
2. 発表標題 音声中の検出語検出におけるWeb検索とWord Vectorを用いたリスコアリング方式
3. 学会等名 情報処理学会研究報告
4. 発表年 2018年

1. 発表者名 金子大祐, 小嶋和徳, 李時旭, 伊藤慶明
2. 発表標題 音声中の検出語検出におけるドキュメントの最尤系列化と上位候補の再照合による検索時間・精度の改善
3. 学会等名 情報処理学会研究報告
4. 発表年 2018年

1. 発表者名 岩崎 瑛太郎, 小原 真人, 小嶋 和徳, 李 時旭, 伊藤 慶明
2. 発表標題 音声中の検索語検出における深層学習の事後確率を用いたクエリの最尤系列化方式,
3. 学会等名 日本音響学会秋季研究発表会
4. 発表年 2018年

1. 発表者名 飯田 英仁, 小嶋 和徳, 李 時旭, 伊藤 慶明
2. 発表標題 遠野物語方言音声の収録とその理解システムの検討
3. 学会等名 日本音響学会秋季研究発表会
4. 発表年 2018年

1. 発表者名 岩崎 瑛太郎, 川村 朋輝, 小嶋 和徳, 李 時旭, 伊藤 慶明
2. 発表標題 音声中の検索語検出における深層学習を用いた最尤系列の最適化による検索精度の向上
3. 学会等名 日本音響学会春季研究発表会
4. 発表年 2018年～2019年

1. 発表者名 金子 大祐, 小嶋 和徳, 李 時旭, 伊藤 慶明
2. 発表標題 音声中の検索語検出におけるドキュメント最尤系列化と複数の機械学習モデルによる検索時間・精度の改善
3. 学会等名 日本音響学会春季研究発表会
4. 発表年 2018年～2019年

1. 発表者名 横田 平志, 小嶋和徳, 眞田尚久, 李時旭, 伊藤慶明
2. 発表標題 音声中の音声検索語検出における平均事後確率ベクトル圧縮方式の検索精度改良
3. 学会等名 情報処理学会第83回全国大会
4. 発表年 2021年

1. 発表者名 西野将弘, 伊藤慶明, 小嶋和徳, 李時旭
2. 発表標題 異種・複数の深層学習モデルを用いた音声中の検索語検出方式の高精度・低メモリ化
3. 学会等名 日本音響学会春季研究発表会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担者	李 時旭 (Lee Shi-wook) (50415642)	国立研究開発法人産業技術総合研究所・情報・人間工学領域・主任研究員 (82626)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分 担 者	松原 雅文 (Matsuhara Masafumi) (70363728)	岩手県立大学・ソフトウェア情報学部・准教授 (21201)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関