

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 12 日現在

機関番号：14603

研究種目：若手研究(B)

研究期間：2012～2013

課題番号：24700172

研究課題名(和文) 少数言語向け多言語音声認識の知識統合フレームワーク

研究課題名(英文) knowledge sources integration framework for multilingual speech recognition for minority language

研究代表者

サクリアニ サクティ (Sakti, Sakriani)

奈良先端科学技術大学院大学・情報科学研究科・助教

研究者番号：30625083

交付決定額(研究期間全体)：(直接経費) 3,400,000円、(間接経費) 1,020,000円

研究成果の概要(和文)：口頭伝承や口承文学など数多くの無形文化は極めてもろく消滅の危機にさらされている。少数民族の多くは言葉や文化の壁による疎外に直面している。音声翻訳は少数民族社会においては自身の言語を守ると共に言葉の壁を克服するツールとして有意義である。本研究では音声翻訳システムの一部である音声認識技術を少数言語向けに発展させる。インドネシアに現存する4つの主要な少数言語からデータ収集を収集できた。音声リソースおよび音声認識ベースラインに基づき、インドネシア少数言語の特性を分析することができた。少数言語の音声コーパスを音声認識ベースラインに適用し、少数言語の音声認識システムを開発することができた。

研究成果の概要(英文)：Various intangible cultural expressions are fragile and easily lost, and many indigenous people face isolation due to language and cultural barriers. Speech translation is significant for indigenous communities to preserve their cultural language and overcome language barriers. This research focuses on one part of speech translation system which developing a speech recognition for minority indigenous languages. We were successfully collect the data from the main minority languages existing in Indonesia, and analyze the characteristics of those languages. Furthermore, by incorporating efficient knowledge source into Indonesian ASR baseline, the ASR of minority languages were successfully been developed.

研究分野：総合領域

科研費の分科・細目：情報学・知覚情報処理、知能ロボティクス

キーワード：音声情報処理 音声認識

1. 研究開始当初の背景

Background

我々の世界は文化の多様性により豊かで生き生きとしたものになる。現在、世界の少数民族の人口は3億5000万人にのぼり、およそ6000種の言語と文化が存在する。一方、グローバル化が進む中、口頭伝承や口承文学などの無形文化は極めてもろく、消滅する可能性が高い。すでにいくつかの国際プロジェクト(2004年~2008年のEPOCHプロジェクト<http://www.epoch-net.org>や2004年~2005年のUNESCO ICT4IDプロジェクトなど)において、情報通信技術(ICT)を文化保存に活用し消滅の危機から救う動きが始動している。これまでに、同じ少数民族社会に属する高齢者と若者のコミュニケーション、あるいは少数民族以外の人とのコミュニケーションを支援する技術が開発されているが、実用にはまだ限界があり、少数民族社会は今もって言葉や文化の壁による阻害に直面している。

音声翻訳は話した言葉を別の言語に音声で翻訳する技術である。この技術により異なる言語を話す人同士のコミュニケーションが可能になり、少数民族社会においては言葉の壁や異文化ギャップを克服するツールとして有用である。特に、人が話した言葉を書き言葉に置き換える自動音声認識は音声翻訳の研究でも重要な技術である。統計的音声認識では大量のモノリンガルコーパスからトレーニングを行う方法が主流であるが、十分なデータが見込めない言語については、過去、異なるデータを融合し多言語音響特性をひとつの構造に組み込む研究が発表された。しかし少数言語の音声認識では、使えるデータが足りないだけでなく、話し言葉として使われることが多く文字として書かれた資料が少ないため、その発話

内容や音韻情報については未知の部分が大きいという問題もある。したがって、本研究のような複雑なシステムを開発し最適性能を得るということは、難しいどころか不可能であるともいわれてきた。

2. 研究の目的

Research purpose

音声翻訳は異なる言語を話す人同士のコミュニケーションを可能にする技術であり、少数民族社会においては自身の言語を守ると共に言葉の壁を克服するツールとして有意義である。本研究では、音声翻訳システムの一部である音声認識技術を少数言語向けに発展させる。提案者は、過去の実績から、音声言語処理におけるグラフィカルモデル技術および教師なしノンパラメトリックベイズ推定法に関する研究知識を獲得している。本研究では異なる知識を融合し多言語音響特性をひとつの構造に組み込むグラフィカルモデリングとベイズ推定法による効果的な知識統合フレームワークを提案する。

3. 研究の方法

Research method

(a) アジアの少数言語について調査を行い、本研究で対象とする言語を確定する。さらにグラフィカルモデルフレームワークとベイズ推定法についての知識を深め、当該技術で最先端の知識を有する世界の研究者と議論を行う。

(b) データ収集: データ収集はアジアの複数の国・地域で行う。

(c) 分析および設計: 少数言語に含まれる知識を分析する。発話内容だけでなく非発話内容も分析の対象とする。その後、多言語音声認識向けの知識統合フレームワーク

を設計・構築する。

(d) システム構築：音声認識タスクの各レベル（フレーム、音素、単語、文レベル）で知識を統合する方法を探索し、システムを構築する。

(e) 性能評価：応用システムにおいてフレームワークの効果を評価する。

4. 研究成果

Research achievement

(1) 研究調査

本プロジェクトにより、音声処理関連の会議に出席する機会を得ることができ、少数言語の音声認識研究に関して多くの研究者と議論を交わすことができた。

(2) 少数言語のデータ収集

本プロジェクトでは、インドネシアに現存する4つの主要な少数言語からデータ収集を始めた。

*ジャワ語：ジャワ島中部で使用

*スンダ語：ジャワ島西部で使用

*バリ語：バリ島で使用

*バタック語：スマトラ島北部で使用

結果、下記少数言語のコーパスを収集することができた。

(a) テキストコーパス

インドネシア少数言語のテキストコーパスは、地元の新聞・雑誌から収集し、ジャワ語、スンダ語、バリ語、バタック語の各少数言語につき2000~3000文を収集できた。うち、1000文を選んでネイティブの話者による校閲を行った。その後、Greedy アルゴリズムを用いて音素バランス付き文を処理し、各言語につき225の音素バランス付き文を生成した。

(b) 音声コーパス

*音素バランス付き音声コーパス

上記の音素バランス付き225文を用い、各

少数言語につき10人（男性5人、女性5人）の話者による音声データの収集を行った。収録はインドネシア国内で実施した。ジャワ語、スンダ語、バリ語、バタック語につきそれぞれ音素バランス付きコーパスを合計2250文収集した。

*対訳音声コーパス

音素バランス付きコーパスの他に、インドネシア語からジャワ語、スンダ語、バリ語、バタック語に翻訳された50文の対訳音声コーパスも収集した。この音声データについては、各言語につき10人（男性5人、女性5人）の話者による発話を収録した。

(3) 音声認識システムベースライン

*インドネシア語音声認識システム

インドネシア語の音声認識ベースラインは、既存のインドネシア語音声データ（ニュースおよび電話会話）を用いて学習した。使用したコーパスは、PT Telekomunikasi Indonesia (R&D TELKOM) のR&D部門がAPT (Asia Pacific Telecommunity) の継続プロジェクトとしてATRと共同開発したものである。このコーパスでは、400人の話者（男性200人、女性200人）の発話が収録されている。話者の発話は、標準インドネシア語のアクセントと3つのネイティブ言語（バタック語、ジャワ語、スンダ語）のアクセントが付けられている。各話者は210文を発話しており、合計84,000発話、80時間の音声が取められている。

(4) 少数言語音声認識システム

上記の音声リソースおよび音声認識ベースラインに基づき、インドネシア少数言語の特性を分析することができた。また、少数言語の音声コーパスを音声認識ベースラインに適用し、少数言語の音声認識システムを開発することができた。データ収集は、事務手続き等が原因で当初の予定よりも長

くかかった。その結果、音声言語処理におけるグラフィカルモデル技術の適応にまで進めることができなかった。現時点では、各インドネシア少数言語に対しGraphemeベースの音声認識を適用している。その認識結果は、各言語においてWER <10%を達成している。今後は、グラフィカルモデルを用いてひとつの多言語認識フレームワークに統合すべく研究を進める。

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)

Publication (underline your name)

[雑誌論文] (計 件)

Journal

[学会発表] (計 3 件)

Conference

(1) Sakriani Sakti, Satoshi Nakamura, “Recent Progress in Developing Grapheme-based Speech Recognition for Indonesian Ethnic Languages: Javanese, Sundanese, Balinese, and Bataks, in Proc. of SLTU 2014, 2014年05月14日, St. Petersburg, Russia.

(2) Sakriani Sakti, Satoshi Nakamura, “Towards Language Preservation: Design and Collection of Graphemically Balanced and Parallel Speech Corpora of Indonesian Ethnic Languages”, in Proc. of Oriental COCOSDA 2013, 2013年11月25日, Delhi, India.

(3) Auliya Sani, Sakriani Sakti, Graham Neubig, Tomoki Toda, Adi Mulyanto, Satoshi Nakamura. “Towards Language

Preservation: Preliminary Collection and Vowel Analysis of Indonesian Ethnic Speech Data”, in Proc. of Oriental COCOSDA 2012, 2012年12月11日, Macau China.

[図書] (計 件)

Book

[産業財産権]

○出願状況 (計 件)

Patent (in application)

名称:

発明者:

権利者:

種類:

番号:

出願年月日:

国内外の別:

○取得状況 (計 件)

Patent (obtained)

名称:

発明者:

権利者:

種類:

番号:

取得年月日:

国内外の別:

[その他]

ホームページ等

Homepage etc.

6. 研究組織

(1) 研究代表者

サクリアニ サクティ (SAKTI, Sakriani)

奈良先端科学技術大学院大学・情報科学

研究科・助教

研究者番号：30625083

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：