

様式 F - 7 - 1

科学研究費助成事業（学術研究助成基金助成金）実施状況報告書（研究実施状況報告書）（平成25年度）

1. 機関番号

1	4	6	0	3
---	---	---	---	---

 2. 研究機関名 奈良先端科学技術大学院大学

3. 研究種目名 基盤研究(C) 4. 補助事業期間 平成24年度～平成26年度

5. 課題番号

2	4	5	0	0	1	1	5
---	---	---	---	---	---	---	---

6. 研究課題名 機械学習および統計モデルに基づく音声対話システムの応答生成の研究

7. 研究代表者

研究者番号	研究代表者名	所属部局名	職名
8 0 3 3 5 4 8 9	カワナミ ヒロミチ 川波 弘道	情報科学研究科	助教

8. 研究分担者

研究者番号	研究分担者名	所属研究機関名・部局名	職名
0 0 2 6 3 4 2 6	シカノ キヨヒロ 鹿野 清宏	情報科学研究科	教授
3 0 3 2 4 9 7 4	サルワタリ ヒロシ 猿渡 洋	情報科学研究科	准教授

9. 研究実績の概要

音声をユーザ入力とした情報案内システムはユーザ負担が少ないことから、万人に優しい情報アクセス方法として期待されている。我々は用例ベースで応答生成を行うアプローチを採用した実環境システムを開発し、子供、大人に対して60%、70%の応答正解率を達成しているが、本課題では更なる応答性能向上のため機械学習によるトピック分類と自動翻訳技術による応答文生成技術に取り組んでいる。

本研究課題ではこれまで、応答生成に先立ちユーザ発話を10数個のトピックに分類する課題を設定し、用例ベースによるトピック分類を有意に卓越する成果を実現した。これは機械学習の入力特徴量として音声認識結果のBOW(Bag-of-Words)を用いることと、複数の識別器の出力結果を再度識別器への入力として用いる Stacked Generalization手法の導入で実現された。昨年度は我々が10年間に及ぶシステム運用で得た大量の、ラベル付けしていないユーザ発話を用いた半教師あり学習による自動的な性能向上の可能性を調査した。加えて、応答生成の前処理である不要入力（雑音や不要発話）棄却性能の向上にも取り組んだ。特徴量として従来の音響的特徴量の音声・雑音モデル尤度に加え、やはりBOWを導入した。その結果、従来手法を超える棄却率を実現した。

これらと並行して応答生成に自動翻訳技術を導入する研究にも取り組んだ。翻訳モデルにより質問文を応答文に変換することは新奇性が高く、複数仮説が得られる自動認識結果が変換元学習データとして利用できるためローコストになるという利点もある。音声認識結果からの応答文生成の結果、書き起こし文で構築した翻訳モデルより適切な応答文が生成されることが分かった。しかしながら従来手法である用例ベースの応答性能には到達しておらず、H25年度は課題点の分析を中心に調査を行った。

10. キーワード

(1) 音声対話システム	(2)	(3)	(4)
(5)	(6)	(7)	(8)

11. 現在までの達成度

(区分)(2) おおむね順調に進展している。

(理由)

機械学習によるトピック分類では、BOW(Bag-of-Words)特徴量とStacked Generalizationによる2段階分類を導入することで大人・子供の実環境発話を用いた実験で有意な改善が見られた。同様の特徴量を用いて、雑音のみならずシステム応答不要な発話も棄却することをめざした、不要入力棄却実験においても従来手法を超える棄却性能を実現し、実用的な手法であることが示された。トピック分類に関しては更なる性能向上を目指してラベル付与されていない大量コーパスを利用した半教師あり学習にも取り組んだ。有意な性能向上を実現するアルゴリズムは検討中であるが、このアプローチの検証は10年以上にわたる実システム運用で得たデータを活用できる我々の研究グループの優位性により実現可能となったものである。

もう一方のアプローチである統計的自動翻訳を用いた応答文生成の研究では、学習データに音声認識結果を用いたことで応答性能は向上したが、従来手法での応答性能には達していなかった。実験結果の分析により、本手法を適用した場合に悪影響を与える・受けるデータの性質が明らかになった。その知見を活かすことで最終年度では本アプローチの性能向上が期待できる。

これらの研究の進捗状況から、H24年度はおおむね順調に研究を進めることができたとして自己評価するものである。

12. 今後の研究の推進方策 等

(今後の推進方策)

最終年度であるH26年度は、大量のラベルなし音声コーパスを活用した半教師ありトピック分類における有効な学習アルゴリズムの検討を行う。

また、本年度後半には本研究課題遂行によりすでに従来手法に対する優位性が示された、BOW(Bag-of-Words)による不要入力棄却処理、トピック分類処理を実システムに組み込み、その総合的な応答性能とユーザの印象を評価対象とした実証実験を行う。

機械翻訳技術による応答文生成のアプローチでは、翻訳結果の分析からノイズを発生すると考えられるデータ、つまり省略、冗長、フィルターを含むデータを除外したデータでの性能を調査する。そしてそのようなスクリーニングと学習データ量との関係を調査することで手法の限界やコストを考慮して本手法の評価を行う。

(次年度使用額が生じた理由と使用計画)

(理由)

既存の音声入出力機材・計算機に経年劣化による不安定動作が見られたため、H26年度(最終年度)後半に計画している音声情報案内システム実証実験のために器材を補充する必要があるが、実験を行う時点での最新の器材を補充しりこととしたため。

(使用計画)

次年度後半期開始時期にマイク、スピーカ、計算機の置き換えを行う。

13.研究発表(平成25年度の研究成果)

(雑誌論文) 計(0)件 うち査読付論文 計(0)件

著者名		論文標題			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
掲載論文のDOI(デジタルオブジェクト識別子)					

(学会発表) 計(2)件 うち招待講演 計(0)件

発表者名		発表標題	
Shunsuke Nakai, Ryoichi Miyazaki, Hiromichi Kawanami, Hiroshi Saruwatari, Satoshi Nakamura		Semi-blind multiple speech extraction under diffuse noise condition for smart posterboard	
学会等名	発表年月日	発表場所	
NCSP '14 (RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing)	2014年02月28日 ~ 2014年03月03日	Waikiki Beach Marriott Resort & Spa, Honolulu, Hawaii, USA	

発表者名		発表標題	
Yuka Hirano, Ryoichi Miyazaki, Hiromichi Kawanami, Hiroshi Saruwatari, Satoshi Nakamura, Tomoya Takatani		Unsupervised Control of Speech Quality Based on Higher-Order Statistics in Musical-Noise Free Blind Speech Extraction	
学会等名	発表年月日	発表場所	
NCSP '14 (RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing)	2014年02月28日 ~ 2014年03月03日	Waikiki Beach Marriott Resort & Spa, Honolulu, Hawaii, USA	

