

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24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 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。すでに我々は用例ベースの実環境システムを開発し70～60%の応答正解率を達成しているが、更なる性能の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を導入した。

H24年度は、機械学習によるトピック分類に2段階分類を導入することで、大人・子供発話とも分類性能を改善した。具体的には15カテゴリーの分類をタスクとした教師あり学習による実験を行った。まず1段目の識別処理として音声認識結果のBOW（Bag-of-words；単語頻度ベクトル）を特徴量としてサポートベクターマシン、最大エントロピー法、pboostの3手法による識別を行い、それらの出力結果を特徴量としたサポートベクターマシンにより2段目の最終的な識別を行う、Stacked Generalization手法を導入した。

加えて、実システムにおいてトピック分類処理の前段階に必要な不要入力（雑音や不要発話）棄却実験も行った。特徴量として従来用いられている音響的特徴、音声・雑音モデルの尤度等に加えここでもBOWを追加した。その結果、従来手法を超える棄却率を実現した。

応答生成に自動翻訳技術を導入する研究では、音声認識結果がシステム入力であることをふまえ、翻訳モデルの翻訳元データとして音声認識を用いた実験を行った。音声認識結果を用いることは翻訳モデル構築コストにも繋がる。音声認識結果は一つの音声データから順位付きで複数の候補仮説が得られるので学習、テストに複数の候補仮説を用いた評価も行った。これらの方法を導入することで、書き起し文による翻訳モデルのものより適切な応答文が生成された。しかしながら、質問文データベースの探索による用例ベースの方式の応答性能にはまだ到達しておらず、誤りの要因分析が必要な状況となっている。

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の研究は、2段階分類を導入することで大人・子供発話とも分類性能の有意な改善が実現できた。実験タスクとしても実用的な15カテゴリーのトピック分類であり、実環境運用により得られた一般のユーザ発話を用いた実験であることから、この手法が実システムで有効であることを示すことができた。

同様の枠組を採用して、実システムデータの不要入力棄却実験も行ったが、これも従来手法を超える棄却率を実現し、実用的な手法であることが示された。

統計的自動翻訳を用いた応答文生成に関しては、書き起し文を変換元とする我々のこれまでの手法と比べて、正しい応答文、適切な表現の応答文が生成される率は向上したが、質問文データベースの探索による用例ベースの方式の応答性能には到達していない。ただし用例ベースの手法では得られない表現の多様性においては利点がある。

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)

H24年度には、それまでに進めてきた教師ありデータによるトピック分類の技術を総合的に比較評価することで、有効な基本的な枠組を獲得することができた。そこで、トピック分類の高精度化に半教師あり学習を導入する。我々が本研究で使用している音声データは、現在に至るまで10年以上にわたって実環境運用している音声対話システムにより得られている。そのうち24年度で利用したデータは人手による書き起しを行った2年強のデータのみである。その以外の大量の音声データを教師なしデータとして用いることで、データがあれば自動的に分類性能向上が可能となる枠組を構築する。

翻訳モデルによる応答文生成のアプローチに関しては、音声認識結果を変換元とした翻訳モデルが応答生成に有効であることが示唆された。用例ベースに基づく既存技術の応答正解率には到達しなかったが、既存手法では実現不可能な「入力の表現に対応して表現を多様化できる」という利点がある。そこで生成された応答内容、表現の不適切な点について要因分析を行い、翻訳モデルの特徴量、モデルパラメータの詳細な検討を行う。

最終年度には、不要入力棄却処理、機械学習による識別モデルによるトピック分類、翻訳モデルによる応答文生成技術を組み込み、実証実験を行う。

(次年度の研究費の使用計画)

該当なし

13.研究発表(平成24年度の研究成果)

〔雑誌論文〕計(2)件 うち査読付論文 計(2)件

著者名		論文標題【掲載確定】			
真嶋温佳, 藤田洋子, トーレスラファエル, 川波弘道, 原直, 松井知子, 猿渡洋, 鹿野清宏		音声情報案内システムにおけるBag-of-Wordsを用いた無効入力棄却			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
情報処理学会論文誌	有	Vol.54 No.2	2 0 1 3	443-451	
掲載論文のDOI(デジタルオブジェクト識別子)					
なし					

著者名		論文標題【掲載確定】			
Rafael Torres, Hiromichi Kawanami, Tomoko Matsui, Hiroshi Saruwatari, Kiyohiro Shikano		Comparison of Methods for Topic Classification of Spoken Inquiries			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
Journal of Information Processing	有	Vol.21, No.2	2 0 1 3	157-167	
掲載論文のDOI(デジタルオブジェクト識別子)					
10.2197/ipsjip.21.157					

〔学会発表〕計(3)件 うち招待講演 計(0)件

発表者名		発表標題【発表確定】	
Kazuma Nishimura, Hiromichi Kawanami, Hiroshi Saruwatari, Kiyohiro Shikano		Response generation based on statistical machine translation for speech-oriented guidance system	
学会等名	発表年月日	発表場所	
Proc. APSIPA Annual Summit and Conference 2012 (APSIPA2012)	2012年12月04日	Hollywood USA	

発表者名	発表標題【発表確定】	
真嶋温佳, トーレス ラファエル, 川波弘道, 原直, 松井知子, 猿渡洋, 鹿野清宏	音声情報案内システムにおけるBag-of-Wordsを用いた無効入力棄却モデルの可搬性の評価	
学会等名	発表年月日	発表場所
日本音響学会春季講演会	2013年03月15日	東京都八王子市

発表者名	発表標題【発表確定】	
吉田 雄太, 久保 慶伍, 川波 弘道, 猿渡 洋, 鹿野 清宏	音声情報案内システム「たけまるくん」の応答誤り分析とその改善	
学会等名	発表年月日	発表場所
日本音響学会春季講演会	2013年03月15日	東京都八王子市

(図書) 計(0)件

著者名	出版社		
書名	発行年	総ページ数	

14. 研究成果による産業財産権の出願・取得状況

(出願) 計(0)件

産業財産権の名称	発明者	権利者	産業財産権の種類、番号	出願年月日	国内・外国の別

(取得) 計(0)件

産業財産権の名称	発明者	権利者	産業財産権の種類、番号	取得年月日	国内・外国の別
				出願年月日	

15.備考

--