

論文 / 著書情報
Article / Book Information

| | |
|-------------------|--|
| 題目(和文) | 残響および騒音のある環境下で頑健な音声認識 |
| Title(English) | Robust Speech Recognition under Reverberant and Noisy Environments |
| 著者(和文) | 太刀岡勇気 |
| Author(English) | Yuki Tachioka |
| 出典(和文) | 学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10885号, 授与年月日:2018年3月26日, 学位の種別:課程博士, 審査員:小林 隆夫,奥村 学,山口 雅浩,金子 寛彦,篠崎 隆宏 |
| Citation(English) | Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第10885号, Conferred date:2018/3/26, Degree Type:Course doctor, Examiner:,,,, |
| 学位種別(和文) | 博士論文 |
| Category(English) | Doctoral Thesis |
| 種別(和文) | 審査の要旨 |
| Type(English) | Exam Summary |

(博士課程)

論文審査の要旨及び審査員

| 報告番号 | 甲第 | 号 | 学位申請者氏名 | 太刀岡 勇気 | | |
|-------------|-----|-------|---------|--------|-------|-----|
| 論文審査 審査員 | | 氏名 | 職名 | | 氏名 | 職名 |
| | 主査 | 小林 隆夫 | 教授 | 審査員 | 篠崎 隆宏 | 准教授 |
| | 審査員 | 奥村 学 | 教授 | | | |
| | | 山口 雅浩 | 教授 | | | |
| 金子 寛彦 | | 教授 | | | | |

論文審査の要旨 (2000 字程度)

本論文は「Robust Speech Recognition under Reverberant and Noisy Environments (残響および騒音のある環境下で頑健な音声認識)」と題し、英文6章からなっている。

第1章「Introduction (序論)」では、研究の背景として実環境で音声認識システムを使用した際に性能が低下する原因として騒音と残響があることを述べ、音声認識の適用先拡大のため騒音・残響がある環境下で音声認識性能を向上させることが研究目的であると述べている。

第2章「Sound source localization methods (音源位置推定手法)」では、複数マイクロホンを使って話者の位置を特定し、この情報を基に騒音の影響を軽減する技術として、音源の方向に関する事前分布を利用した音源方向推定方法と、マイクロホンの配置誤差・残響による理論的なマイク間の到来時間差からのずれを補正する方法を提案し、実験を通して提案手法の有効性を示している。

第3章「Front-end techniques for robust automatic speech recognition (ASR) (頑健な音声認識のための前処理手法)」では、性能を向上させるための音声認識の前処理として、騒音・残響の影響を軽減し音声信号成分を復元する音声強調技術と、騒音・残響下で音声信号が存在する区間のみを切り出す音声区間検出技術について詳細に検討している。音声強調技術に関しては、マイクロホンを1つ使う方法として残響時間を自動的に推定し残響を除去する方法、複数のマイクロホンを使う方法として到来方向に基づくバイナリマスクと独立ベクトル分析を統合する方法、マルチチャンネル非負値行列因子分解に基づく音声強調手法に対して性能を安定させる初期値を与える方法をそれぞれ提案し、これらの音声強調手法を適用することにより音声認識性能が大幅に向上することを実験的に示している。一方音声区間検出に関しては、モデルパラメータ推定に密度比推定を使うことにより頑健性を向上させる手法を提案している。この他にも、システム学習時と使用時における環境のミスマッチに起因する音声認識性能の低下を抑えるいくつかの前処理手法について言及している。

第4章「Back-end techniques for robust ASR (頑健な音声認識のための後処理手法)」では、音声認識システムを騒音・残響に対して頑健にする方法に関して述べている。第2章及び第3章で提案した手法により騒音を抑圧できるとはいえ、騒音の影響を完全に除去することはできないことから、音声認識処理そのものを騒音に対して頑健にする必要がある。ここでは、特徴量の線形判別分析と音声認識処理に用いられる音響モデル・言語モデルに識別学習を導入し、騒音環境下で音声認識性能が向上することを示している。さらに、音響モデルのモデル削減法と識別学習を組み合わせる方法、識別学習の目的関数を拡張して元のシステムと異なる仮説を出力するような複数のシステムを組み合わせて用いる枠組み、音声強調に伴って混入した音声のひずみの影響による認識性能劣化を低減する方法を提案している。

第5章「Development of ASR systems for realistic noisy and reverberant environments (騒音・残響環境下で動作する音声認識システムの開発)」では、騒音・残響に頑健な音声認識タスクを対象とする様々なチャレンジ(国際コンペティション)を通して、提案手法の有効性を明らかにしている。家庭環境での騒音を対象とした第2回 CHIME チャレンジ、屋外環境の騒音を対象とした第3回及び第4回 CHIME チャレンジ、種々の残響環境を対象とした REVERB チャレンジにおいてそれぞれ、上述の提案手法を用いて構築した音声認識システムの性能を示し、特に第2回 CHIME チャレンジ、REVERB チャレンジでは参加者中トップの成績を収め、提案法の有効性が確認されたと述べている。

第6章「Conclusion (結論)」では、本論文で得られた成果をまとめ、今後の課題を述べている。

以上を要するに、本論文では騒音・残響がある環境で音声認識性能を向上させることを目的として、音源の位置推定法、音声強調法(騒音抑圧・残響除去)、音声認識の各種モデルの頑健性を向上させる識別的手法を提案し、公開チャレンジにより提案法の有効性を確認したものあり、工学上ならびに工業上貢献するところが大きい。よって本論文は博士(工学)の学位論文として価値あるものと認められる。

注意: 「論文審査の要旨及び審査員」は、東工大リサーチポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。