

論文 / 著書情報
Article / Book Information

題目(和文)	携帯端末における音声認識のための効率的な誤り訂正
Title(English)	Efficient Error Correction for Mobile Speech Recognition
著者(和文)	梁原
Author(English)	Yuan Liang
出典(和文)	学位:博士(学術), 学位授与機関:東京工業大学, 報告番号:甲第10018号, 授与年月日:2015年9月25日, 学位の種別:課程博士, 審査員:篠田 浩一,亀井 宏行,徳永 健伸,室田 真男,村田 剛志,藤井 敦
Citation(English)	Degree:, Conferring organization: Tokyo Institute of Technology, Report number:甲第10018号, Conferred date:2015/9/25, 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

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		梁 原	
論文審査 審査員		氏 名	職 名		氏 名	職 名	
	主査	篠田 浩一	教授	審査員	村田 剛志	准教授	
	審査員	亀井 宏行	教授		藤井 敦	准教授	
		徳永 健伸	教授				
		室田 真男	教授				

論文審査の要旨 (2000 字程度)

この論文は、“Efficient Error Correction for Mobile Speech Recognition”と題し、英文7章から成っている。

第1章「Introduction」では、研究の背景について述べた上で本論文の構成を示している。まず、IT技術の発展によりスマートフォンなどの携帯情報端末が急速に普及していること、それに伴い、それらの携帯情報端末での音声による情報入力需要が高まっていることを述べた上で、本論文は、携帯情報端末での音声認識のための効率的な誤り訂正インタフェースとそのためのアルゴリズムをその研究対象とすることを述べている。特に、メールなどの比較的長い文章を入力する場面を想定し、そこでのユーザの負荷を減らすことを目的とすることを述べている。

第2章「Automatic Speech Recognition (ASR) System」では、現在主流の統計的音声認識技術を概観した上で、そのための誤り訂正技術の理解に必要な項目についても説明している。特徴抽出、音響モデル、言語モデル、サーチの各要素技術について述べた後、連続音声認識の評価基準と、そこにおける音声認識誤りの分類について述べている。

第3章「Error Correction for Speech Recognition: A Review」では、携帯情報端末における音声認識のための誤り訂正について、その従来技術を概観し、その評価方法について述べている。特に、認識器が出力する Word Confusion Network (WCN) から作成された候補単語リストを提示し、そこからユーザに選択をさせる方法が主流であり、候補単語リストの上位に正解単語が位置するようにすることで、ユーザの負荷を減らすことができること、また、そのために誤り単語区間の音響尤度やユーザによる過去の誤り訂正入力の情報の利用が有効であることが述べられている。さらに、ユーザの入力ストローク数がユーザの負荷を測る基準としてしばしば用いられることを述べている。

第4章「Gesture Based Error Correction Interface」では、携帯情報端末における音声認識のための、よりユーザの負荷が少ない誤り訂正インタフェースを提案している。このインタフェースでは、ユーザは誤り箇所を誤りの種類毎に異なるストロークで指定することを述べている。また、長い文章を入力する場合には必ずしも誤り箇所の指定と同時にその誤りの訂正を行う必要がないことから、次章で説明するアルゴリズムを実現するために、ある誤り箇所の誤り訂正は文章の先の新たな誤り箇所の指定をした後に可能になることを述べている。

第5章「Long Context Match (LCM) Method for Generating Candidate Word List」では、候補単語リストにおける正解単語の順位を向上させるためのアルゴリズムについて述べている。まず、誤り訂正を誤り箇所の指定よりも遅らせて行うことにより、誤り箇所の前の単語コンテキストだけでなく、その後の単語コンテキストも利用可能になることを述べ、前後の単語コンテキストと大規模テキストコーパス中の単語列とのマッチングにより誤り訂正を行う方法 (LCM) を提案している。さらに、LCM と音響尤度や WCN を用いる方法との組み合わせも提案している。

第6章「Experiments」では、第4・5章で述べたインタフェースとアルゴリズムの評価について述べている。音声データとして日本語話し言葉コーパス Corpus of Spontaneous Japanese (CSJ) を、誤り訂正用テキストコーパスとして Google n-gram を用いた評価において、従来の WCN を用いる方法や誤り箇所の前の単語コンテキストのみを用いる方法と比べ、提案法での候補単語リスト中1位の単語の正解率が7.0%以上高いことを述べている。また、ユーザの誤り訂正を計算機で模擬する評価実験で、想定される入力ストローク数が従来法に比べ12%減少したことを述べている。

第7章「Conclusions and Future Work」では本論文で得られた成果をまとめ、将来の研究における課題について述べている。

以上で述べたように、本論文では、携帯情報端末における音声認識のための誤り訂正インタフェースとそのためのアルゴリズムを提案し、その有効性を評価実験を通して確認している。本論文で得られた成果は、音声言語処理の発展に貢献するとともに、ヒューマンインタフェース分野との学際領域において学術上寄与するところが大きい。よって本論文は博士(学術)の学位論文として十分価値があるものと認める。

注意:「論文審査の要旨及び審査員」は、東工大リサーチリポジトリ(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。