

## 論文内容の要旨

博士論文題目 音響空間可視化手法を応用した効率的な音声コーパス  
構築フレームワーク

氏名 奈木野 豪秀

### (論文内容の要旨)

音声認識アプリケーションの開発において高い性能を実現するには、使用するタスクに特化した音響モデルが必須となる。タスクに特化した音響モデルを作成するために、タスクの音響的特性を考慮した大量の音声データの収集が行われることが多い。しかし、音声データの収集は膨大なコストを必要とし、音声認識アプリケーションの開発コスト全体を圧迫している。さらに、データ収集システムの開発や構築音声データベースの設計には、音声認識の専門家の経験則に頼るところが大きく、音声認識アプリケーション開発のボトルネックとなっている。

本論文は、音響空間の可視化手法、既存音声コーパスの目的タスクに対する再利用性の判定手法、収録話者の選択による音声コーパス構築の低コスト化の手法を提案している。まず、音響空間の可視化手法COSMOS法を考案して、音響空間の可視化に成功している。このCOSMOS法により、タスク間、タスク内の音響的変動、実環境で発生する雑音と音声との違いを直感的に把握できることを実証している。従来の可視化手法と比較して、データの分布の分離度においても優位性を示している。

既存音声コーパスのタスクに対する再利用性の判定において、従来法では既存音声コーパス群から音響的に特徴の近い音声データを選択することはできるが、選択後に作成された音響モデルの性能を保証することが困難であった。COSMOS法による再利用性判定手法を用いることで、目的タスクと既存タスクとの可視空間上の分布の重なり具合から、直感的に再利用性を把握することが可能となることを実証している。さらに、収集対象の候補話者の少量音声データから可視化手法を用いて音声認識性能向上に寄与する話者を予備選択して、選択された話者の音声データを収集する手法を提案している。従来の無作為に話者を選択する手順と比較して、大幅なコスト削減が可能となることを実証している。

これらの成果は、実際に音声認識の開発現場で利用されており、音声認識アプリケーションの開発コストの大幅な削減効果を達成している。

氏名	奈木野 豪秀
----	--------

(論文審査結果の要旨)

音声認識アプリケーションの開発において高い性能を実現するには、使用するタスクに特化した大量の音声データによる音響モデルが必須となる。しかし、音声データの収集は膨大なコストを必要とし、音声認識アプリケーションの開発コスト全体を圧迫している。

本論文では、音響空間の可視化手法COSMOS法を考案して音響空間を可視化に成功している。このCOSMOS法は、従来の可視化手法と比較して、データの分布の分離度においてもはるかに優れていることを、タスク間、タスク内の音響的変動などに応用して、直感的に把握できることを実証し、学会でも高く評価されている。

既存音声コーパスのタスクに対する再利用性の判定においても、COSMOS法による再利用性判定手法を用いることで、目的タスクと既存タスクとの可視空間上の分布の重なり具合から、直感的に再利用性を把握することが可能となることも実証し、大幅な音声データ収集コストの削減が可能になることを示した。

これらの成果は、学术论文2件、権威のある音声の国際会議 ICASSP や ICSLP に8件採録されて発表しており、国際的にも高く評価されている。さらに、実際に音声認識の開発現場で利用しており、音声認識アプリケーションの開発のコスト削減効果を達成している。

音声認識の分野を深く理解した技術者としての音声認識技術によるビジネス展開だけでなく、国際的な舞台での活躍もできる研究者に育っており、今後の活躍が期待できます。

平成20年7月29日に開催した公聴会の結果も参考にして、本博士論文の審査を行い、本論文は、博士論文(工学)として十分な価値があるものと判断した。