

報告番号	※甲	第	号
------	----	---	---

主論文の要旨

論文題目 声帯振動の物理特徴に基づくストレス環境
下での音声の検出

氏名 姚 瀟

論文内容の要旨

ストレスを誘発するような環境での音声に対しては、自動音声認識 (ASR) システムの精度は低下する。そのような環境の例としては、騒々しい環境、複数のタスクを実行している場合、疲労がある場合、感情的な場合、体調不良な場合、作業量が過度である場合、などがあげられる。ストレス音声に関する研究は、これらの状況下での音声認識システムの頑健性を維持あるいは改善することに貢献できる。本論文は物理モデルに基づいてストレス音声を分類することを主目的として論じる。ストレスに対して、話者は生理的变化を起こし、ストレス状況に対応したり、適応したりする。この変化により、喉頭と声道の空気動力学の変化が生まれ、ストレス発話が生成される。したがって、このような発音過程を表現する気流パターンの物理モデルが必要である。本論文では、物理モデルを用いたストレス音声の検出方法を検討する。以下に、本論文での目的を示す。

- I. ストレス発話での声帯の物理特性を究明し、ストレス検出に用いることが可能なパラメータを推定する (研究 1)。
- II. ストレス発話の声帯と声道の物理特性を究明し、識別に必要なパラメータを提案する (研究 2)。
- III. 喉頭室と仮声道の空気動力特性をモデル化し、ストレス音声を検出するためのパラメータを推定する (研究 3)。

IV. 異なる識別方法を用いてストレス音声を識別し、結果を比較する (研究 4).

研究 1 では、物理モデルに基づいてストレス音声を識別する手法を提案する。音声生成過程を表す物理モデルにより、喉頭内の気流パターンをモデル化する。モデルのパラメータには発音する時の声帯の特性が反映されている。Two-mass モデルを利用し、実の音声から、声帯筋肉の緊張度、声帯粘性の喪失度、肺からの喉頭下圧力をそれぞれ反映する物理パラメータの推定手法を提案する。これらのパラメータの組み合わせが、正常音声とストレス音声の識別に有効な特徴量として提案する。

研究 2 では、声帯と声道両方の特性を検討する。実の音声の物理モデルに基づいて声帯と声道の物理パラメータを推定する方法を提案する。物理モデルでは、two-mass モデルと four-tube を組み合わせて発音過程を模擬する。実の音声データからモデルのパラメータ（ばね定数、声道長、声道断面積）を学習する。推定された物理パラメータと実際の分類表現に対して、異なる評価用のコスト関数を提案する。

研究 3 では、喉頭室と仮声帯の空気動力学変化パターンを用いて正常音声とストレス音声を分類することを検討する。喉頭室と、心理的ストレス下の喉頭室内の気流パラメータを含むように two-mass モデルを改良する。Two-mass モデルに基づいて実の音声から喉頭室の有効面積パラメータを推定する。評価実験は推定されたパラメータが正常音声とストレス音声の識別に有効であることを示す。

最後の研究では、我々は異なる識別器を用いてストレス音声の識別性能を比較する。物理的特徴量に対し、線形分類器、混合ガウスモデル、SVM それぞれを用いてモデル化する。

提案法の物理特徴量と従来の特徴量に対して比較を行った。実験結果より、ストレス音声の識別性能が改善されることが分かった。声道を表現するパラメータは母音依存での識別に有効であり、声帯パラメータの方がストレス識別に対してより有効である。さらに、喉頭室と仮声帯の気流変化特徴を表現するパラメータを加えることで、よりよい識別性能が得られた。

本研究の応用について、心理抑圧音声の検出より電話詐欺の防止できる。それ以外に、運転手の音声のストレスの検出より運転安全を確保することもできる。この技術は広い領域に応用され、感情音声認識の表現を改善することも可能と考える。異なるストレス環境に対応して、提案法と他の方法を組み合わせて、

音声認識の精度を上げることもできると考える.