

論文 / 著書情報  
Article / Book Information

題目(和文)	
Title(English)	Comprehensive Feature Optimization for Robust Speaker Recognition Across Diverse Datasets
著者(和文)	CHAUHANNeha
Author(English)	Neha Chauhan
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12859号, 授与年月日:2024年9月20日, 学位の種別:課程博士, 審査員:一色 剛,高橋 篤司,本村 真人,原 祐子,佐々木 広
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12859号, Conferred date:2024/9/20, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

(博士課程)  
Doctoral Program

## 論文要旨

THESIS SUMMARY

系・コース : Department of, Graduate major in	Information and Communications	系 コース	申請学位 (専攻分野) : Academic Degree Requested	博士 Doctor of	( )	Engineering
学生氏名 : Student's Name	Chauhan Neha		審査員主査 : Chief Examiner	Tsuyoshi Isshiki		

要旨 (英文 300 語程度)

Thesis Summary (approx.300 English Words )

Speaker recognition systems have gained significant attention in recent years due to their wide-ranging applications in security and authentication. This thesis presents a comprehensive exploration of advanced techniques to enhance the performance of speaker recognition models. The research is divided into two major parts, each emphasizing distinct methodologies to improve the accuracy and efficiency of speaker recognition systems and other part is reducing computational timing.

The first part focuses on a novel fusion technique that combines 18 different speech features, including mel frequency cepstral coefficient (MFCC), linear predictive coding (LPC), perceptual linear prediction (PLP), root mean square (RMS), centroid, and entropy features. By incorporating their respective delta ( $\Delta$ ) and delta-delta ( $\Delta\Delta$ ) feature vectors, this approach demonstrates a substantial enhancement in speaker identification accuracy and a notable reduction in equal error rates (EER) for speaker verification across diverse datasets. Notably, the experiments conducted on the NIST-2008, voxforge, ELSDSR, VCTK, and voxceleb1 speech corpora showcase remarkable improvements, with the highest achieved speaker identification accuracy reaching 100% and the lowest EER plummeting to 0%.

The second part of the research delves into a comprehensive study on speaker recognition, emphasizing three pivotal strategies: feature-level fusion, dimension reduction using Principal Component Analysis (PCA) and Independent Component Analysis (ICA), and feature optimization employing genetic algorithm (GA) and marine predators algorithm (MPA). By implementing a combination of PCA with GA and MPA, this approach significantly enhances the performance of speaker recognition systems, particularly evident in various datasets with diverse noise levels and speaker counts. Notably, the experimental results showcase a notable improvement in speaker identification accuracy, reaching up to 95.2%, and a considerable reduction in SV EER down to 1.8%, notably boosting the computational efficiency, particularly for the challenging voxceleb1 dataset.

This thesis provides key insights and methodologies for creating robust speaker recognition systems adaptable to diverse speech datasets, contributing to the advancement of reliable biometric security measures and authentication protocols.

備考 : 論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note : Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意 : 論文要旨は、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。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).