氏　　名	GALAJIT, Kasorn
学位の種類	博士（情報科学）
学位記番号	博情第 454 号
学位授与年月日	令和 3 年 9 月 24 日
論文題目	Singular Spectrum Analysis-based Auditory Information Hiding with CNN-based Parameter Estimation for Speech Security
論文審査委員	主査　鵜木　祐史　北陸先端科学技術大学院大学　教授
	赤木　正人　同　教授
	党　建武　同　教授
	吉高　淳夫　同　准教授
	KARNJANA Jessada　NECTEC　研究員
	AIMANEE Pakinee　タマサート大学 SIIT　准教授

論文の内容の要旨

Speech signals are adopted in various forms and many social applications in a cyber-physical system (CPS), such as voice command, voice activation, and voice recognition. However, high-end speech editing software, such as voice conversion techniques and speech synthesis software, makes anyone easily fabricate and alter speech signals. These misused of this technology create risk in the security of speech technology and lead to social problems according to the increasing of unauthenticated speeches. These unauthenticated speeches can be used for criminal purposes such as theft or fraud in any systems in CPS. The attacks of unauthenticated speech signals, such as tampered speech, spoofed speech, and modified speech are considered an emerging threat. Thus, it is necessary to provide security of speech signals. Cryptography is a classical method that provides security by concealing speech signals from being tampered with and modified. However, cryptography does not detect tampering and modification in speech signals. Auditory information hiding (AIH) is one of the solutions to provide speech security by creating a secret channel and detecting tampering.

This research aims to provide security for speech signals in two objectives. The first objective is security in terms of protecting the genuineness of the speech signal. If attackers try to modify or change the speech signal, AIH can be used to protect its genuineness by tampering detection. One crucial property of information hiding is that the hidden information should be difficult to remove from the watermarked signal, and if there are attacks performed on the watermarked signal, the hidden information should reflect that change. The second objective is to protect the secret communication of the speech signal. AIH can be used to build the secret channel, and the transformation is used to secure the secret data on the secret channel.

Based on literature reviews, several information hiding techniques have been previously developed, and the singular spectrum analysis (SSA)-based AIH showed its strength in robustness due to the invariance of the singular spectrum. Moreover, SSA-based AIH could be designed to gain

semi-fragile property (robust against non-malicious attacks but fragile to malicious attacks) by properly selecting part of the singular spectrum to be modified. The possibility of semi-fragile in SSA-based AIH motivates to construct a scheme for tampering detection. In addition, we deployed the convolutional neural network (CNN) method for parameter estimation instead of the differential evolution-based method adopted in the original SSA-based AIH.

For the first objective, the experimental results showed that the proposed scheme could locate tampered areas correctly, and it could also predict the types and degrees of tampering roughly. CNN-based parameter estimation could significantly reduce computational time, and the scheme is entirely blind because the estimation could be used to suggest the parameters in both embedding and extraction processes. However, the tampering detection accuracy needs to be improved since the proposed scheme fragile to MP4 and robust to echo adding.

For the second objective, we cooperate transformation techniques with our SSA-based AIH to construct the secret and secured channel. The experimental results show that SSA-based AIH cooperated with Arnold transformation technique can provide the secret and secured channel. Only the authorized person with the correct key can access data at each level.

Index Terms: Singular spectrum analysis, SSA-based information hiding, CNN-based parameter estimation, tampering detection, speech security

論文審査の結果の要旨

近年，情報通信技術（ICT）の急速な発達やインフラの整備により，インターネット上でのマルチメディア情報（テキスト，音楽・音声，静止・動画像など）の利用が盛んになっている．特に，サイバーフィジカル空間における音声情報通信の利用は，スマートフォンの普及や AI スピーカの登場，音声コンテンツのアプリ開発の増加とともに，この数年で急激な伸びを示している．このような急激な需要拡大に対して，音声情報を安心・安全に利用するための技術革新や法整備は相当な遅れをとっており，音声コンテンツの違法コピー・違法配信といった社会問題だけでなく，音声改ざんや音声プライバシー侵害，音声なりすましといった問題も招いている．そのため，サイバーフィジカル空間において，ディジタル表現された音声メディア情報を安心・安全に利用するために，ディジタル音信号への汎用性の高い情報ハイディング技術の基盤を確立する必要がある．特に，情報秘匿が難しい音声信号に対し，どのような聴覚特性を考慮した情報ハイディングが有効であるか，どのような音声情報表現（符号化・情報圧縮）が有効であるかを明らかにするとともに，知覚不可能性，頑健性，情報秘匿性を備えた音声情報ハイディング技術基盤を整備する．

本研究では，(1) ディジタル音声信号の真正性を担保する方法と(2) 音声信号の秘密通信を保護する方法を実現するために，データ駆動型の分析手法である特異スペクトル分析に基づいた音声情報ハイディング法を確立した．この方法は，特異スペクトル分析に基づく音情報ハイディング手法において，知覚不可能性と頑健性のトレードオフを考慮したうえで，畳み込みニューラルネットワークによる秘匿情報を正確に検出する方法を実現した．一つの大きな成果として，リサンプリングや再量子化，

音声合成といった意味のある信号変換に対しては頑健で，悪意ある攻撃には脆弱な情報ハイディングを実現した．その結果，音声改ざん検出といった音声メディアのセキュリティを大幅に高めることに成功した．もう一つの大きな成果は，Arnold 変換を利用した音声情報ハイディングを実現したことで，音声信号の秘密通信のための安心で安全な情報伝送系の確立に成功した．

以上，本論文は，サイバーフィジカル空間におけるディジタル音声信号のセキュリティを高めるための一つの革新的技術を提供した．本技術は，音声情報ハイディングとして，音声改ざん防止や音声なりすまし防止といった重要課題に対して応用範囲が広く，学術的に貢献するところも大きい．よって博士（情報科学）の学位論文として十分価値あるものと認めた．