

令和 4 年 6 月 20 日現在

機関番号：13903

研究種目：基盤研究(A)（一般）

研究期間：2018～2021

課題番号：18H04112

研究課題名（和文）PRISM：選択的情報マスキングを可能とする次世代音声プライバシー保護技術

研究課題名（英文）PRISM: Speech privacy preservation based on selecting masking

研究代表者

徳田 恵一（Tokuda, Keiichi）

名古屋工業大学・工学（系）研究科（研究院）・教授

研究者番号：20217483

交付決定額（研究期間全体）：（直接経費） 33,900,000円

研究成果の概要（和文）：本研究では、音声に含まれる個人性と発話内容などのプライバシー情報を選択的に保護することを可能とする次世代音声プライバシー保護技術（PRISM: PRIVacy Selecting Masking）を確立することを目指し、実空間における音声プライバシー保護については合成音声を用いたマスキングによる手法を、サイバースペースにおける音声プライバシー保護については深層学習による機械学習による手法を提案し、それぞれ有効性を確認した。

研究成果の学術的意義や社会的意義

スマートスピーカーなどの音声を入力とするシステムの普及により、音声に含まれるプライバシー情報が収集されるといったプライバシーの侵害が懸念されている。音声には、個人の特定だけでなく、年齢や性別、発話内容など、様々なプライバシー情報が含まれているが、本研究で確立した選択的音声プライバシー保護技術によれば、場所や状況、発話内容等に応じ、それぞれに適した形で音声プライバシー情報を保護することが可能となり、今後のネットワーク化された情報社会において、安心して音声インタフェースを利用することが可能となる。

研究成果の概要（英文）：This research aims to establish a next-generation voice privacy protection technology (PRISM: PRIVacy Selecting Masking) that can selectively protect privacy information such as personal identity and speech content contained in speech. We have proposed a synthetic voice-based masking method for voice privacy protection in real space and a deep learning machine learning method for voice privacy protection in cyberspace, and have confirmed the effectiveness of each method.

研究分野：音声情報処理

キーワード：音声プライバシー マスキング 音声合成 音声情報処理

1. 研究開始当初の背景

スマートフォン、スマートスピーカーなどの音声を入力とするシステムの普及により、音声に含まれるプライバシー情報が収集されるといったプライバシーの侵害が懸念されている。年齢や性別、発話内容など、音声には様々なプライバシー情報が含まれているが、保護したいプライバシー情報は個人によるところが強く、また、同じ人でも場所や状況、発話内容に依存してどのプライバシー情報を保護したいと考えるかが異なることもある。これまでもいくつかの音声プライバシー保護技術が提案されてきているが、特定のプライバシー情報を対象としており、人々が本当に保護したいプライバシー情報に対して過不足が生じている。ソーシャルメディア上の音声データから個人の音声合成システムを作ること容易になりつつある現在、音声に含まれる話者情報、および、保護を行うことが適切な特徴を匿名化やマスキングする技術が求められており、喫緊の対策が必要である。音声に含まれる保護がセンシティブな特徴は、発話内容・話者性の他に、話者の性別、黒人・白人・エスニック等の人種・環境に起因する方言、話者が保有する音声障害情報などが考えられる。理想的には、これらのセンシティブな特徴が、ユーザーの希望や法制度の定めに応じて、選択的にマスキング処理できるようになることが望ましいと考えられ、それを支える基盤技術の確立が求められている。

2. 研究の目的

本研究では、音声に含まれるプライバシー情報を分離・変換可能な形で音声をモデル化することで、音声モデリングによる統一的な枠組みによって選択的にプライバシー情報を保護可能にする、次世代音声プライバシー保護技術 (PRISM: PRiVacy Selecting Masking) を確立することを目的とする。本研究では、音声の匿名化法の研究、とりわけ、音声に含まれる話者に起因する特徴を匿名化する手法「話者匿名化」を中心に研究を進める。本テーマは非常に新しい研究トピックであり、研究開始当時、音声分野を幅広く見渡しても、どの様に話者匿名化を実現できるのか未だ明確に定義されていないのが現状であった。そこで、話者匿名化手法の提案・改善のみならず、話者匿名化音声の評価方法の確立も目的とした。また、実空間における音声プライバシーについては、空間に放射された音声を加工することはできないため、新しいマスキング技術を確立することを目的とする。

3. 研究の方法

本研究では、収録済みの音声データを対象としたプライバシー保護技術(「サイバースペースにおける音声プライバシー保護技術」と実空間における音声を対象としたプライバシー保護技術(「実空間における音声プライバシー保護技術」)の2つの技術の構築に取り組む(図1)。サイバースペースにおける音声プライバシー保護技術では、音声そのものを加工可能である状況を想定し、保存される音声データを変換・保存することによってプライバシー情報を保護する技術を対象とする。音声から読み取ることが可能なプライバシー情報としては年齢や性別、感情などのノンバーバルな情報と、特定の固有名詞や方言、使用言語(日本語、英語など)

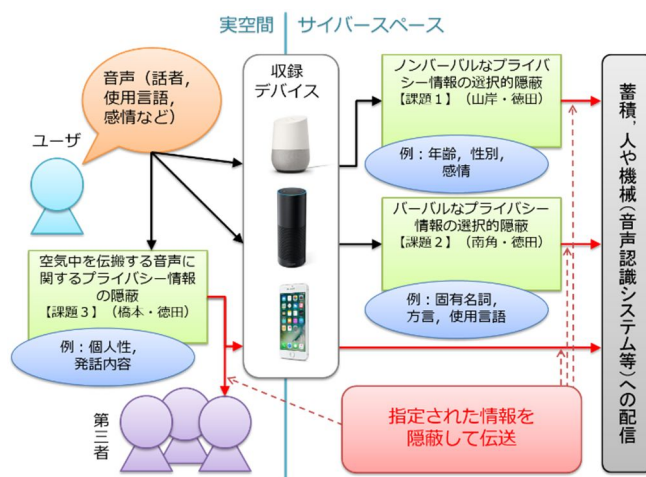


図1 実空間およびサイバースペースにおける音声プライバシーの保護

といったバーバルな情報があり、それぞれを対象としたプライバシー保護技術の確立に取り組む。具体的には、音声の自然性や音声から知覚可能な年代や性別といった属性情報を保ったまま、話者に起因する音声の個人性のみを変換する話者匿名化変換手法を深層学習による機械学習により実現する枠組みについて検討する。また同時に、話者匿名化手法を相互に比較できる様、利用する音声データベース、評価セット、評価手順を規定し、その評価方法自身についても検討する。一方、実空間における音声プライバシー保護技術では、音声そのものを加工することはできず、マスキング用の信号を加算的に生成することによってプライバシー情報を保護する技術を対象とする。特に、音声合成の枠組みを利用し、ユーザーと同じ声質で異なる発話内容の音声

を複数同時に生成することで効率的に発話内容を隠蔽する手法について検討する。

4. 研究成果

(1) 話者匿名化手法の提案

音声の自然性や音声から知覚可能な年代や性別といった属性情報を保ったまま、話者に起因する音声の個人性のみを変える話者匿名化変換手法を提案した[1]。これは、音声を抑揚、音素情報、xベクトルという話者性を表す潜在変数ベクトルの3つの情報に分解し、xベクトルのみを、本人以外のK人の話者から抽出し平均化したxベクトルと置き換えることで匿名化する手法である。音声波形を再合成するモジュールにはニューラルボコーダを利用し、高品質な音声生成を可能にした。英語話者の音声データを利用した実験から、xベクトルの空間において、匿名化を行うことで、話者認識システムおよび人間の聴覚上の話者識別性能が有意に下がることを確認した。

(2) 話者匿名化手法の改良と応用

前述の話者匿名化手法の有用性は確認されているが、音素情報の抽出には音声認識システムを利用しており、それ故、音素ラベルや発音辞書などの言語依存資源が教師ありデータとして必要であり、システム構築にはコストがかかっていた。そこでこのモジュールを、wav2vec 2.0やHuBERT等の音声波形から事前自己学習したモデルに置き換えた話者匿名化手法を提案した[2]。またその手法が学習データに含まれていない言語の音声データに対して適用可能である事も示した。また、類似システムにおいて、話者情報ではなく、逆に、発話内容をマスク処理する事が可能かも検討した。具体的には、まず音声信号を、発話内容を表した局所潜在変数系列と、話者情報を表したグローバル潜在変数に分解し、再度音声波形を再合成するネットワークを構築し、そして、発話内容を現した局所潜在変数系列の一部を、バブル雑音から抽出した別の潜在変数に置き換える事で発話内容のマスク処理を行う手法を提案した[3][4]。本手法を使う事で、ピープ音等の特殊音を使わずに、かつ、もとの話者性を保ったまま、発話の一部をマスクする事が可能になる。

(3) 実空間における音声プライバシー

実空間における音声プライバシー保護技術の確立については、主に発話内容を対象とし、これらの情報を隠蔽するマスク信号の生成技術について検討を進めた[5]。利用者の合成音声を用いたマスク信号の生成に取り組み、主観評価実験により、利用者の声質の合成音声を用いることで、ホワイトノイズやピンクノイズ、さらには他人の声質の合成音声に対して、より高いマスク効果が得られることを示した。実際の公共空間あるいはそれと同等の環境において、利用者の立場でのプライバシー感の評価やマスク信号により周囲に迷惑をかけていると感じるか、第三者がマスク音をどのように感じるかなどを含んだ評価実験を行い、その結果に基づいてより高度で快適なマスク信号生成技術としてまとめた。

(4) Voice Privacy Challenge

音声のプライバシー保護に関して、分野を牽引し、研究を加速させるため、国際的なチャレンジを運営した。仏 Avignon 大、仏 Eurecom 研究所、仏 Inria 研究所と協力して"Voice Privacy Challenge"を実施した。話者匿名化手法を相互に比較できる様、利用する音声データベース、評価セット、評価手順を規定した。話者匿名化の評価には、本人かどうかを判定する話者照合性能指標に加え、話者匿名化された音声再識別不可能であること、そして、匿名化された音声の他の用途に有用さを計測し、比較する必要があると結論づけた。実際、10数の大学・企業・研究組織が提案した話者匿名化手法をこの指標に基づき相互評価する事を行い、国際会議 Interspeech 2020 および Speaker Odyssey 2020 におけるスペシャルセッションを開催した。また国際ジャーナル誌 Computer Speech & Language において音声プライバシーの特集号も企画した。VoicePrivacy 2020 の成果や知見をまとめたジャーナル論文[6]も発表した。

< 引用文献 >

- Fang, F., Wang, X., Yamagishi, J., Echizen, I., Todisco, M., Evans, N. and Bonastre, J.-F.: Speaker Anonymization Using X-vector and Neural Waveform Models, Proc. 10th ISCA Speech Synthesis Workshop, pp. 155-160 (online), DOI: 10.21437/SSW.2019-28 (2019).
- Xiaoxiao Miao, Xin Wang, Erica Cooper, Junichi Yamagishi, Natalia Tomashenko, "Language-Independent Speaker Anonymization Approach using Self-Supervised Pre-Trained Models" Proc. ISCA Speaker Odyssey 2022
- Jennifer Williams, Junichi Yamagishi, Paul-Gauthier Noe, Cassia Valentini Botinhao, Jean-Francois Bonastre, "Revisiting Speech Content Privacy" ISCA Security and Privacy in Speech Communication 2021 (1st SPSC Symposium)

Jennifer Williams, Jason Fong, Erica Cooper, Junichi Yamagishi, "Exploring Disentanglement with Multilingual and Monolingual VQ-VAE" Speech Synthesis Workshop 2021 (SSW11)

次井貴浩, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, "公共空間におけるスピーチプライバシー保護を目的とした合成音声によるサウンドマスキングの検討," 音声研究会, vol. 119, no. 321, SP2019-38, pp. 55-60, 東京, 日本, 2019年12月.

Natalia Tomashenko, Xin Wang, Emmanuel Vincent, Jose Patino, Brij Mohan Lal Srivastava, Paul-Gauthier Noé, Andreas Nautsch, Nicholas Evans, Junichi Yamagishi, Benjamin O' Brien, Anaïs Chanclu, Jean-François Bonastre, Massimiliano Todisco, Mohamed Maouche, "The VoicePrivacy 2020 Challenge: Results and findings," Computer Speech & Language, Volume 74, 2022

<https://www.voiceprivacychallenge.org>

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 1件／うちオープンアクセス 1件）

1. 著者名 Xin Wang, Shinji Takaki, Junichi Yamagishi, Simon King, Keiichi Tokuda	4. 巻 28
2. 論文標題 A Vector Quantized Variational Autoencoder (VQ-VAE) Autoregressive Neural F0 Model for Statistical Parametric Speech Synthesis	5. 発行年 2019年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech and Language Processing	6. 最初と最後の頁 157-170
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TASLP.2019.2950099	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Xin Wang, Shinji Takaki, Junichi Yamagishi	4. 巻 28
2. 論文標題 Neural source-filter waveform models for statistical parametric speech synthesis	5. 発行年 2019年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech and Language Processing	6. 最初と最後の頁 402-415
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TASLP.2019.2956145	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Yi Zhao, Shinji Takaki, Hieu-Thi Luong, Junichi Yamagishi, Daisuke Saito, Nobuaki Minematsu	4. 巻 6
2. 論文標題 Wasserstein GAN and Waveform Loss-based Acoustic Model Training for Multi-speaker Text-to-Speech Synthesis Systems Using a WaveNet Neural Vocoder	5. 発行年 2018年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 60478-60488
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ACCESS.2018.2872060	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計69件（うち招待講演 2件／うち国際学会 35件）

1. 発表者名 藤本崇人, 橋本佳, 南角吉彦, 徳田恵一
2. 発表標題 学習時と合成時の一貫性を考慮したVAEに基づく自己回帰型sequence-to-sequence音声合成
3. 学会等名 日本音響学会2021年春季研究発表会
4. 発表年 2021年

1. 発表者名 角谷健太, 吉村建慶, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 隠れセミマルコフモデルに基づく構造化アテンションを用いたSequence-to-Sequence音声合成
3. 学会等名 日本音響学会2021年春季研究発表会
4. 発表年 2021年

1. 発表者名 法野行哉, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 周期・非周期成分の分離に基づくニューラルボコーダによる音声波形のモデル化の検討
3. 学会等名 日本音響学会2021年春季研究発表会
4. 発表年 2021年

1. 発表者名 岩田康平, 高木信二, 橋本佳, 南角吉彦, 徳田恵一
2. 発表標題 勾配ブースティング決定木を用いた音声合成手法の検討
3. 学会等名 日本音響学会2021年春季研究発表会
4. 発表年 2021年

1. 発表者名 平光啓祐, 橋本佳, 徳田恵一, 南角吉彦
2. 発表標題 深層学習に基づく音声合成における顔画像を用いた話者適応
3. 学会等名 第18回情報学ワークショップ
4. 発表年 2020年

1. 発表者名 車田智哉, 木下耕介, 吉村建慶, 橋本佳, 南角吉彦, 徳田恵一
2. 発表標題 生成モデルの構造を組み込んだ系列変分オートエンコーダに基づく話者認識
3. 学会等名 第18回情報学ワークショップ
4. 発表年 2020年

1. 発表者名 西村愛理, 藤本崇人, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 出力遅延を考慮したアテンション機構に基づくリアルタイム声質変換
3. 学会等名 第18回情報学ワークショップ
4. 発表年 2020年

1. 発表者名 久野宏彰, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 音声合成における特徴的な発話スタイルの転移学習
3. 学会等名 第18回情報学ワークショップ
4. 発表年 2020年

1. 発表者名 成田哲郎, 吉村建慶, 橋本佳, 南角吉彦, 徳田恵一
2. 発表標題 ニューラルボコーダを用いた音声符号化手法の検討
3. 学会等名 第18回情報学ワークショップ
4. 発表年 2020年

1. 発表者名 大谷眞史, 佐藤優介, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 音声合成における敵対的生成ネットワークを用いた複数言語・複数話者モデリングの検討
3. 学会等名 第18回情報学ワークショップ
4. 発表年 2020年

1. 発表者名 佐々木一匡, 吉村建慶, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 大規模音楽データを活用した汎用WaveNetボコーダ構成法の検討
3. 学会等名 第18回情報学ワークショップ
4. 発表年 2020年

1. 発表者名 厚地俊哉, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 音声プライバシー保護のためのノンパラレル声質変換による話者匿名化の検討
3. 学会等名 第18回情報学ワークショップ
4. 発表年 2020年

1. 発表者名 岩田康平, 高木信二, 橋本佳, 南角吉彦, 徳田恵一
2. 発表標題 勾配ブースティング決定木を用いた高速な音声合成手法の検討
3. 学会等名 第18回情報学ワークショップ
4. 発表年 2020年

1. 発表者名 前川遼太郎, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 深層学習に基づく楽器音合成における音響モデルの比較検討
3. 学会等名 第18回情報学ワークショップ
4. 発表年 2020年

1. 発表者名 Yukiya Hono, Kazuna Tsuboi, Kei Sawada, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, and Keiichi Tokuda
2. 発表標題 Hierarchical Multi-Grained Generative Model for Expressive Speech Synthesis
3. 学会等名 Interspeech 2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Brij Mohan Lal Srivastava, Natalia Tomashenko, Xin Wang, Emmanuel Vincent, Junichi Yamagishi, Mohamed Maouche, Aurélien Bellet, Marc Tommasi
2. 発表標題 Design Choices for X-Vector Based Speaker Anonymization
3. 学会等名 Interspeech 2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Natalia Tomashenko, Brij Mohan Lal Srivastava, Xin Wang, Emmanuel Vincent, Andreas Nautsch, Junichi Yamagishi, Nicholas Evans, Jose Patino, Jean-François Bonastre, Paul-Gauthier No?, Massimiliano Todisco
2. 発表標題 Introducing the VoicePrivacy Initiative
3. 学会等名 Interspeech 2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Andreas Nautsch, Jose Patino, N. Tomashenko, Junichi Yamagishi, Paul-Gauthier No?, Jean-Fran?ois Bonastre, Massimiliano Todisco, Nicholas Evans
2. 発表標題 The Privacy ZEBRA: Zero Evidence Biometric Recognition Assessment
3. 学会等名 Interspeech 2020 (国際学会)
4. 発表年 2020年

1. 発表者名 藤本崇人, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 感情音声合成のためのDirichlet VAE
3. 学会等名 日本音響学会2020年秋季研究発表会
4. 発表年 2020年

1. 発表者名 法野行哉, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 DNNに基づく音声ボコーダにおける周期・非周期成分のモデル化の検討
3. 学会等名 日本音響学会2020年秋季研究発表会
4. 発表年 2020年

1. 発表者名 大谷眞史, 佐藤優介, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 音声合成における敵対的生成ネットワークを用いた複数言語・複数話者モデリング
3. 学会等名 日本音響学会2020年秋季研究発表会
4. 発表年 2020年

1. 発表者名 Takato Fujimoto, Shinji Takaki, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, and Keiichi Tokuda
2. 発表標題 Semi-supervised learning based on hierarchical generative models for end-to-end speech synthesis
3. 学会等名 2020 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4. 発表年 2020年

1. 発表者名 村田舜馬, 藤本崇人, 法野行哉, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 楽譜時間情報を用いたアテンション機構に基づく歌声合成の検討
3. 学会等名 日本音響学会2019年秋季研究発表会
4. 発表年 2019年

1. 発表者名 Fuming Fang, Xin Wang, Junichi Yamagishi, Isao Echizen
2. 発表標題 Audiovisual speaker conversion: jointly and simultaneously transforming facial expression and acoustic characteristics
3. 学会等名 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

1. 発表者名 Lauri Juvela, Bajibabu Bollepalli, Junichi Yamagishi, Paavo Alku
2. 発表標題 Waveform generation for text-to-speech synthesis using pitch-synchronous multi-scale generative adversarial networks
3. 学会等名 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

1. 発表者名 Yusuke Yasuda, Xin Wang, Shinji Takaki, Junichi Yamagishi
2. 発表標題 Investigation of enhanced Tacotron text-to-speech synthesis systems with self-attention for pitch accent language
3. 学会等名 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

1. 発表者名 Xin Wang, Shinji Takaki, Junichi Yamagishi
2. 発表標題 Neural source-filter-based waveform model for statistical parametric speech synthesis
3. 学会等名 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

1. 発表者名 Shinji Takaki, Toru Nakashika, Xin Wang, Junichi Yamagishi
2. 発表標題 STFT spectral loss for training a neural speech waveform model
3. 学会等名 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

1. 発表者名 大浦圭一郎, 中村和寛, 橋本佳, 南角吉彦, 徳田恵一
2. 発表標題 周期・非周期信号を用いたDNNに基づくリアルタイム音声ボコーダ
3. 学会等名 情報処理学会研究報告
4. 発表年 2019年

1. 発表者名 大浦圭一郎, 高木信二, 中村和寛, 橋本佳, 南角吉彦, 徳田恵一
2. 発表標題 周期・非周期信号を用いた敵対的生成ネットワークに基づくリアルタイム音声ボコーダ
3. 学会等名 日本音響学会2019年秋季研究発表会
4. 発表年 2019年

1. 発表者名 Keiichi Tokuda
2. 発表標題 Statistical approach to speech synthesis: past, present and future
3. 学会等名 Interspeech 2019 (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Lauri Juvela, Bajibabu Bollepalli, Junichi Yamagishi, Paavo Alku
2. 発表標題 GELP: GAN-Excited Liner Prediction for Speech Synthesis from Mel-spectrogram
3. 学会等名 Interspeech 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 Mingyang Zhang, Xin Wang, Fuming Fang, Haizhou Li, Junichi Yamagishi
2. 発表標題 Joint training framework for text-to-speech and voice conversion using multi-source Tacotron and WaveNet
3. 学会等名 Interspeech 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 Keiichiro Oura, Kazuhiro Nakamura, Kei Hashimoto, Yoshihiko Nankaku, Keiichi Tokuda
2. 発表標題 Deep neural network based real-time speech vocoder with periodic and aperiodic inputs
3. 学会等名 10th ISCA Speech Synthesis Workshop (SSW10) (国際学会)
4. 発表年 2019年

1. 発表者名 Takato Fujimoto, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, Keiichi Tokuda
2. 発表標題 Impacts of input linguistic feature representation on Japanese end-to-end speech synthesis
3. 学会等名 10th ISCA Speech Synthesis Workshop (SSW10) (国際学会)
4. 発表年 2019年

1. 発表者名 Motoki Shimada, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, Keiichi Tokuda
2. 発表標題 Low computational cost speech synthesis based on deep neural networks using hidden semi-Markov model structures
3. 学会等名 10th ISCA Speech Synthesis Workshop (SSW10) (国際学会)
4. 発表年 2019年

1. 発表者名 Fuming Fang, Xin Wang, Junichi Yamagishi, Isao Echizen, Massimiliano Todisco, Nicholas Evans, Jean-Francois Bonastre
2. 発表標題 Speaker Anonymization Using X-vector and Neural Waveform Models
3. 学会等名 10th ISCA Speech Synthesis Workshop (SSW10) (国際学会)
4. 発表年 2019年

1. 発表者名 Yusuke Yasuda, Xin Wang, Junichi Yamagishi
2. 発表標題 Initial investigation of encoder-decoder end-to-end TTS using marginalization of monotonic hard alignments
3. 学会等名 10th ISCA Speech Synthesis Workshop (SSW10) (国際学会)
4. 発表年 2019年

1. 発表者名 Xin Wang, Junichi Yamagishi
2. 発表標題 Neural Harmonic-plus-Noise Waveform Model with Trainable Maximum Voice Frequency for Text-to-Speech Synthesis
3. 学会等名 10th ISCA Speech Synthesis Workshop (SSW10) (国際学会)
4. 発表年 2019年

1. 発表者名 徳田 恵一
2. 発表標題 統計的音声合成の進展と展望
3. 学会等名 音声研究会 (招待講演)
4. 発表年 2019年

1. 発表者名 次井貴浩, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 公共空間におけるスピーチプライバシー保護を目的とした合成音声によるサウンドマスキングの検討
3. 学会等名 音声研究会
4. 発表年 2019年

1. 発表者名 和田蒼汰, 法野行哉, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 歌声合成におけるニューラルボコーダの比較検討
3. 学会等名 音声研究会
4. 発表年 2019年

1. 発表者名 Hieu-Thi Luong, Junichi Yamagishi
2. 発表標題 BOOTSTRAPPING NON-PARALLEL VOICE CONVERSION FROM SPEAKER-ADAPTIVE TEXT-TO-SPEECH
3. 学会等名 IEEE ASRU 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 藤本崇人, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 End-to-End音声合成のための階層化生成モデルに基づく半教師あり学習
3. 学会等名 日本音響学会2020年春季研究発表会
4. 発表年 2020年

1. 発表者名 Koki Senda, Yukiya Hono, Kei Sawada, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, and Keiichi Tokuda
2. 発表標題 Singing Voice Conversion Using Posted Waveform Data on Music Social Media
3. 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2018) (国際学会)
4. 発表年 2018年

1 . 発表者名 Yukiya Hono, Shumma Murata, Kazuhiro Nakamura, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, and Keiichi Tokuda
2 . 発表標題 Recent Development of the DNN-based Singing Voice Synthesis System -- Sinsy
3 . 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2018) (国際学会)
4 . 発表年 2018年

1 . 発表者名 Takato Fujimoto, Takenori Yoshimura, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, and Keiichi Tokuda
2 . 発表標題 Speech Synthesis Using WaveNet Vocoder Based on Periodic/Aperiodic Decomposition
3 . 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2018) (国際学会)
4 . 発表年 2018年

1 . 発表者名 Kento Nakao, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, and Keiichi Tokuda
2 . 発表標題 Speaker Adaptation for Speech Synthesis Based on Deep Neural Networks Using Hidden Semi-Markov Model Structures
3 . 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2018) (国際学会)
4 . 発表年 2018年

1 . 発表者名 Takenori Yoshimura, Natsumi Koike, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, and Keiichi Tokuda
2 . 発表標題 Discriminative feature extraction based on sequential variational autoencoder for speaker recognition
3 . 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2018) (国際学会)
4 . 発表年 2018年

1. 発表者名 Kei Sawada, Takenori Yoshimura, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, and Keiichi Tokuda
2. 発表標題 The NITech text-to-speech system for the Blizzard Challenge 2018
3. 学会等名 Blizzard Challenge 2018 Workshop (国際学会)
4. 発表年 2018年

1. 発表者名 中尾健人, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 時間構造を考慮したニューラルネットワークに基づく音声合成における話者適応の検討
3. 学会等名 音声研究会
4. 発表年 2018年

1. 発表者名 吉村建慶, 小池なつみ, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 話者認識のためのSequential VAEに基づく特徴抽出の検討
3. 学会等名 日本音響学会2018年秋季研究発表会
4. 発表年 2018年

1. 発表者名 次井貴浩, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 公共空間におけるスマートデバイスを用いた音声入力のためのサウンドマスキングに関する検討
3. 学会等名 日本音響学会2018年秋季研究発表会
4. 発表年 2018年

1. 発表者名 藤本崇人, 吉村建慶, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 周期・非周期成分の分離に基づくWaveNetボコーダを用いた音声合成
3. 学会等名 日本音響学会2018年秋季研究発表会
4. 発表年 2018年

1. 発表者名 法野行哉, 村田舜馬, 中村和寛, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 Deep Neural Networkに基づく歌声合成システム -- Sinsy
3. 学会等名 日本音響学会2018年秋季研究発表会
4. 発表年 2018年

1. 発表者名 沢田慶, 吉村建慶, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 Blizzard Challenge 2018のためのNITechテキスト音声合成システム
3. 学会等名 日本音響学会2018年秋季研究発表会
4. 発表年 2018年

1. 発表者名 島田基樹, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 隠れセミマルコフモデルの構造を用いたDNNに基づく音声合成における計算量削減手法の検討
3. 学会等名 日本音響学会2019年春季研究発表会
4. 発表年 2019年

1. 発表者名 藤本崇人, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 周期・非周期信号から駆動するディープニューラルネットワークに基づく音声ボコーダ
3. 学会等名 日本音響学会2019年春季研究発表会
4. 発表年 2019年

1. 発表者名 法野行哉, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 敵対的生成ネットワークを用いた歌声合成の検討
3. 学会等名 日本音響学会2019年春季研究発表会
4. 発表年 2019年

1. 発表者名 角谷健太, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一
2. 発表標題 DNNに基づく感情音声合成のための敵対的学習の検討
3. 学会等名 日本音響学会2019年春季研究発表会
4. 発表年 2019年

1. 発表者名 Shinji Takaki, Toru Nakashika, Xin Wang, Junichi Yamagishi
2. 発表標題 STFT spectral loss for training a neural speech waveform model
3. 学会等名 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

1. 発表者名 Xin Wang, Shinji Takaki, Junichi Yamagishi
2. 発表標題 Neural source-filter-based waveform model for statistical parametric speech synthesis
3. 学会等名 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

1. 発表者名 Lauri Juvola, Bajibabu Bollepalli, Junichi Yamagishi, Paavo Alku
2. 発表標題 Waveform generation for text-to-speech synthesis using pitch-synchronous multi-scale generative adversarial networks
3. 学会等名 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

1. 発表者名 Fuming Fang, Xin Wang, Junichi Yamagishi, Isao Echizen
2. 発表標題 Audiovisual speaker conversion: jointly and simultaneously transforming facial expression and acoustic characteristics
3. 学会等名 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

1. 発表者名 Fuming Fang, Junichi Yamagishi, Isao Echizen, Md Sahidullah, Tomi Kinnunen
2. 発表標題 Transforming acoustic characteristics to deceive playback spoofing countermeasures of speaker verification systems
3. 学会等名 WIFS2018: IEEE International Workshop on Information Forensics and Security (国際学会)
4. 発表年 2018年

1 . 発表者名 Hieu-Thi Luong, Junichi Yamagishi
2 . 発表標題 Scaling and bias codes for modeling speaker-adaptive DNN-based speech synthesis systems
3 . 学会等名 2018 IEEE Workshop on Spoken Language Technology (SLT 2018) (国際学会)
4 . 発表年 2018年

1 . 発表者名 Yusuke Yasuda, Xin Wang, Shinji Takaki, Junichi Yamagishi
2 . 発表標題 Investigation of enhanced Tacotron text-to-speech synthesis systems with self-attention for pitch accent language
3 . 学会等名 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4 . 発表年 2019年

1 . 発表者名 Huy H. Nguyen, Junichi Yamagishi, Isao Echizen
2 . 発表標題 Capsule-Forensics: Using Capsule Networks to Detect Forged Images and Videos
3 . 学会等名 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4 . 発表年 2019年

1 . 発表者名 Yukiya Hono, Kazuhiro Nakamura, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, and Keiichi Tokuda
2 . 発表標題 Singing voice synthesis based on generative adversarial networks
3 . 学会等名 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (国際学会)
4 . 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

VoicePrivacy initiative https://www.voiceprivacychallenge.org/ Baseline Recipe for VoicePrivacy Challenge 2020 https://github.com/Voice-Privacy-Challenge/Voice-Privacy-Challenge-2020 VoicePrivacy initiative https://www.voiceprivacychallenge.org/ Baseline Recipe for VoicePrivacy Challenge 2020 https://github.com/Voice-Privacy-Challenge/Voice-Privacy-Challenge-2020
--

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	山岸 順一 (Yamagishi Junishi) (70709352)	国立情報学研究所・コンテンツ科学研究系・教授 (62615)	
研究分担者	南角 吉彦 (Nankaku Yoshihiko) (80397497)	名古屋工業大学・工学(系)研究科(研究院)・准教授 (13903)	
研究分担者	橋本 佳 (Kei Hashimoto) (10635907)	名古屋工業大学・工学(系)研究科(研究院)・准教授 (13903)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計1件

国際研究集会 Voice Privacy Challenge (https://www.voiceprivacychallenge.org/)	開催年 2020年～2020年
---	--------------------

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------

フランス	国立情報学自動制御研究所 (INRIA)	Eurecom研究所	アビニオン大学	
英国	エジンバラ大学	シェフィールド大学		