

令和 2 年 6 月 29 日現在

機関番号：62615

研究種目：若手研究(A)

研究期間：2017～2019

課題番号：17H04687

研究課題名(和文) 劣悪環境に頑健な声のデジタルクローン技術の確立とその応用

研究課題名(英文) Robust voice cloning technologies in noisy environments and its applications

研究代表者

山岸 順一 (Yamagishi, Junichi)

国立情報学研究所・コンテンツ科学研究系・教授

研究者番号：70709352

交付決定額(研究期間全体)：(直接経費) 16,700,000円

研究成果の概要(和文)：話者適応は音声合成を応用した声のデジタルクローン技術である。本研究は音声合成用途以外の劣環境で収録された音声を対象とすべく、必要な要素技術を新たに提案し、データベースも構築した。まず低品質音声と元の高品質音声が対となったパラレルデータベースDR-VCTKを構築した。そしてテキストデータが付随しない音声信号からでも手軽に声のデジタルクローン出来るように、Multi-modal architectureという新たなニューラルネットを提案した。さらに劣環境下の音声を元に学習した話者エンコーダを組み込んだ新たなニューラルネットワークも提案し、教師なし話者適応を劣環境下の音声からでも行えることを示した。

研究成果の学術的意義や社会的意義

音声合成用音響モデルの学習は、通常、スタジオ収録した高品質音声のみを対象にする。それゆえ、雑音・反響音を含む音声もしくは低品質収録器材により収録された音声に基づき音声合成を行うことは容易ではなく、研究理論に至っては全く構築されていないと言って良い状況であった。本研究は既存技術の制約を取り払い、劣悪条件や正解ラベルがないと言った環境においても、声のデジタルクローンを可能にする。それゆえ、音声合成および話者適応技術を理論的により熟成させたという学術的意義を持つ。また、音声合成および話者適応技術の応用先が爆発的に増えると予想され、社会的意義も大きい。

研究成果の概要(英文)：The "digital voice cloning technology" is an application of speaker adaptation in speech synthesis. In this study, we proposed new elemental technologies and constructed a database for speech recorded in inferior environments other than speech synthesis. First of all, a parallel database "DR-VCTK" was constructed in which low quality voice and original high quality voice are paired. We proposed a new neural network structure called "Multi-modal architecture" to enable the digital cloning of voices from voice data without text data. In addition, a new neural network structure incorporating a speaker encoder that uses speech recorded in a poor environment as training data is proposed, and it is shown that unsupervised speaker adaptation can be performed even from low-quality speech recorded in a non-ideal environment.

研究分野：音声情報処理

キーワード：音声情報処理 音声合成 深層学習 話者適応 音声強調

## 1. 研究開始当初の背景

音声合成は、テキストなどの入力から人工的に音声を生成する技術である。音声信号処理、自然言語処理、音声特徴量と言語特徴量の対応関係を学習・予測する機械学習、聴覚心理学等に及ぶ横断的学術分野であり、従来の応用例には、カーナビゲーション、視覚障害者のスクリーンリーダ、音声障害者の会話補助器、音声対話システム等がある。音響特徴量と言語特徴量の対応関係の学習法には、隠れマルコフモデル、ディープニューラルネットワークなど様々な機械学習が用いられる。音声および機械学習の著名な国際学会(ICASSP, Interspeechなど)で多くの発表があり、Google DeepMindが大量データおよび複雑なニューラルネットワークで肉声に近い合成音声を実現するなど、合成音声の品質は日々着実に向上している。

我々はこの音声合成手法を活用し、たった数分程度の少量の音声データで目標の声や発話様式を再現する、いわば「声のデジタルクローン技術」を研究してきた。この技術を利用し、騒音下で音声を聞きやすくする、音声翻訳システムをパーソナライズする、音声障害者の個人用音声合成システムを構築する事などに応用可能であることを示した。

## 2. 研究の目的

本研究の目的は、音声合成用途以外の劣環境で収録された音声を新たにデジタルクロンの対象とすべく、必要な要素技術を先駆的に生み出す事である。とりわけ、耐雑音・耐反響性を向上させ、高価な音声収録器材を不要とする、ディープラーニングによる話者適応、及び、教師なし話者適応手法を新たに提案することである。既存音声合成技術の制約を取り払うことで、ニーズが多く存在するコンテンツ業界や放送・エンターテインメント業界などへの利用可能になる様に技術を高度化させる事が狙いである。

## 3. 研究の方法

### 課題1

音声合成の品質劣化は、雑音や反響音だけでなく、収録の際に利用される収録機器によっても起きる。大量生産されているスマートフォンに搭載されているマイクロフォンは、通常収録スタジオで利用される高価なマイクロフォンに比べ、収録可能周波数領域や感度が大きく異なり、いわゆる「安っぽい」音になる。またマイクロフォンと口の間の距離もより離れたものになることから、近接効果が薄れ、低周波数域を正しく収録することができない。最初の課題は、この様な安価だが手軽な機材により収録された音声の特徴を、あたかも高価なマイクでスタジオ収録したかの様な高品質音声へ自動変換する技術の検討および必要なデータベースの構築である。

### 課題2

通常、音声合成では、音素等のバランスや頻度を考慮して人工的に作成された「音素バランス文」を読み上げた音声を利用する。しかしながら、故人の声をデジタルクローンにより再現するという様な応用を考えた場合、新たに読み上げ音声を収録するという選択肢は無く、収録済みの会話や対話音声といった必ずしもテキストデータが付随しない音声データにより音声合成システムを構築する必要がある。そこで、次なる課題は、音声デジタルクローン技術のさらに利便性を向上させる教師なし適応についての検討である。この教師無し適応は任意の音声データのみから話者適応を行う技術である。

### 課題3

新話者適応技術の有効性を裏付けるため、課題1と2を統合した技術を開発し、新たな応用例において評価を行う。

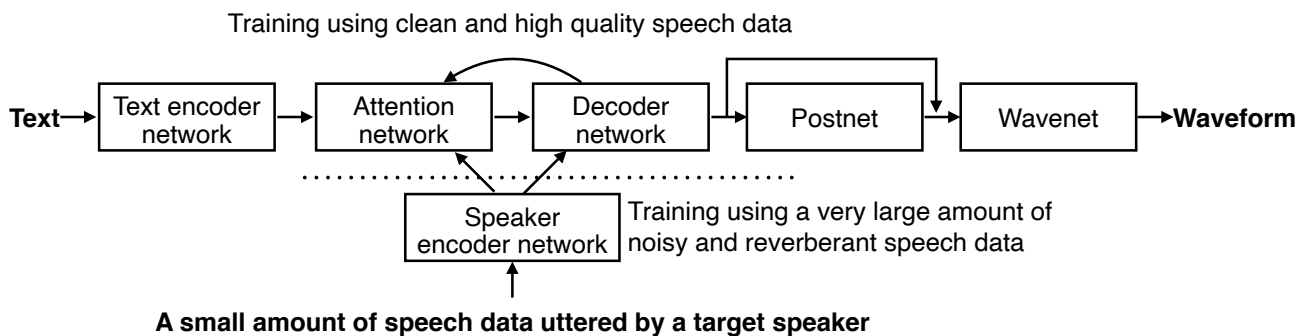
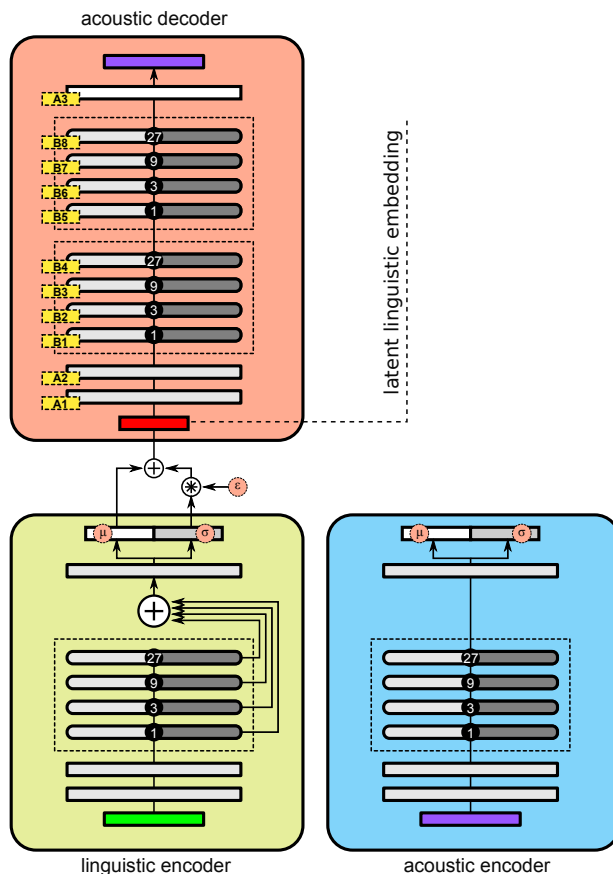
## 4. 研究成果

平成29年度は、課題1に取り組み、安価だが手軽な機材により収録された音声の特徴を、あたかも高価なマイクでスタジオ収録したかの様な高品質音声へ自動変換することを目標に、スタジオ収録された高品質な音声をスピーカで再生し、種々の安価だが手軽に利用可

能な機器により再収録することで、低品質音声と元の高品質音声が対となったパラレルデータベース「DR-VCTK」を構築した。また、このデータベースをもとに、安価な機器により収録された音声により高品質な音声へ変換するニューラルネットワークを学習し、さらに音声合成システムを構築するというも行い、研究発表やジャーナル論文の投稿を行った。

平成30年度は、課題2に取り組み、テキストデータが付随しない音声データからでも手軽に声のデジタルクローン出来るように、Multi-modal architectureという右図に示した新たなニューラルネットワーク構造を提案し、この構造を利用することによりテキストデータが付随しない音声のみでも話者適応が行えることを示した。これらの成果を国際学会Interspeech 2018に投稿し、発表を行った。さらに、合成音声の品質はボコーダという音響特徴量を音声波形信号に変換する技術により大きく制約されてしまうことから、このボコーダの改良も鋭意行った。Neural source-filter modelなどの新たなニューラル波形モデルを提案し、国際会議ICASSP2019等にて複数の論文発表を行った。

令和元年度は、課題1と課題2を同時に解決することを目的に、劣環境で収録された音声を学習データとして利用した「話者エンコーダ (Speaker encoder)」を組み込んだ新たなニューラルネットワーク音声合成方式(下図)を新たに提案し、教師なし話者適応を劣環境で収録した音声からでも実現できることを実験的に示した。実験結果は国際会議ICASSP2020に投稿し、論文採択された。



そして、合成音声の品質もさらに向上させるため、ニューラル波形モデルの改良も鋭意行い、Neural Harmonic-plus-Noise Waveform Modelなど改良版を提案し、ジャーナル論文の投稿を行い、採択された。さらに、声のデジタルクローン技術を実データに対しても適用し、その有効性を検討すると同時に、声のデジタルクローン技術の新たな応用例も模索することも行った。具体的には、プロの囃家の落語実演を収録した音声に対して、前述の話者エンコーダを組み込んだニューラルネットワーク音声合成方式およびニューラル波形モデルを適用することで、囃家が使い分ける様々な役を再現する落語音声合成が実現できることを示した。実験結果は国際会議SSW2019に投稿し、論文採択された。

## 5. 主な発表論文等

〔雑誌論文〕 計30件（うち査読付論文 30件 / うち国際共著 14件 / うちオープンアクセス 29件）

1. 著者名 Airaksinen Manu, Juvela Lauri, Bollepalli Bajibabu, Junichi Yamagishi, Alku Paavo,	4. 巻 26(9)
2. 論文標題 A comparison between STRAIGHT, glottal, and sinusoidal vocoding in statistical parametric speech synthesis	5. 発行年 2018年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech and Language Processing	6. 最初と最後の頁 1658-1670
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/TASLP.2018.2835720">https://doi.org/10.1109/TASLP.2018.2835720</a>	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 該当する
1. 著者名 Igor Jauk, Jaime Lorenzo-Trueba, Junichi Yamagishi, Antonio Bonafonte	4. 巻 -
2. 論文標題 Expressive Speech Synthesis Using Sentiment Embeddings	5. 発行年 2018年
3. 雑誌名 Proc. Interspeech 2018	6. 最初と最後の頁 3062--3066
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.21437/Interspeech.2018-2467">http://dx.doi.org/10.21437/Interspeech.2018-2467</a>	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 該当する
1. 著者名 Lauri Juvela, Vassilis Tsiaras, Bajibabu Bollepalli, Manu Airaksinen, Junichi Yamagishi, Paavo Alku	4. 巻 -
2. 論文標題 Speaker-independent Raw Waveform Model for Glottal Excitation	5. 発行年 2018年
3. 雑誌名 Proc. Interspeech 2018	6. 最初と最後の頁 2012--2016
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.21437/Interspeech.2018-1635">http://dx.doi.org/10.21437/Interspeech.2018-1635</a>	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 該当する
1. 著者名 Hieu-Thi Luong, Junichi Yamagishi	4. 巻 -
2. 論文標題 Multimodal Speech Synthesis Architecture for Unsupervised Speaker Adaptation	5. 発行年 2018年
3. 雑誌名 Proc. Interspeech 2018	6. 最初と最後の頁 2494--2498
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.21437/Interspeech.2018-1791">http://dx.doi.org/10.21437/Interspeech.2018-1791</a>	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 -

1. 著者名 Yi Zhao, Shinji Takaki, Hieu-Thi Luong, Junichi Yamagishi, Daisuke Saito, Nobuaki Minematsu	4. 巻 6(1)
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) <a href="https://doi.org/10.1109/ACCESS.2018.2872060">https://doi.org/10.1109/ACCESS.2018.2872060</a>	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 -

1. 著者名 Hieu-Thi Luong, Junichi Yamagishi	4. 巻 -
2. 論文標題 Scaling and bias codes for modeling speaker-adaptive DNN-based speech synthesis systems	5. 発行年 2018年
3. 雑誌名 2018 IEEE Spoken Language Technology Workshop (SLT)	6. 最初と最後の頁 610-617
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/SLT.2018.8639659">https://doi.org/10.1109/SLT.2018.8639659</a>	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 -

1. 著者名 Toru Nakashika, Shinji Takaki, Junichi Yamagishi,	4. 巻 27(2)
2. 論文標題 Complex-Valued Restricted Boltzmann Machine for Speaker-Dependent Speech Parameterization From Complex Spectra	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 244-254
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/TASLP.2018.2877465">https://doi.org/10.1109/TASLP.2018.2877465</a>	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 -

1. 著者名 Takaki Shinji, Nakashika Toru, Wang Xin, Yamagishi Junichi	4. 巻 -
2. 論文標題 STFT Spectral Loss for Training a Neural Speech Waveform Model	5. 発行年 2019年
3. 雑誌名 2019 IEEE International Conference on Acoustics, Speech and Signal Processing	6. 最初と最後の頁 7065-7069
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/ICASSP.2019.8683791">https://doi.org/10.1109/ICASSP.2019.8683791</a>	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 -

1. 著者名 Xin Wang, Shinji Takaki, Junichi Yamagishi	4. 巻 -
2. 論文標題 Neural source-filter-based waveform model for statistical parametric speech synthesis	5. 発行年 2019年
3. 雑誌名 2019 IEEE International Conference on Acoustics, Speech and Signal Processing	6. 最初と最後の頁 5916-5920
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/ICASSP.2019.8682298">https://doi.org/10.1109/ICASSP.2019.8682298</a>	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 -

1. 著者名 Yusuke Yasuda, Xin Wang, Shinji Takaki, Junichi Yamagishi	4. 巻 -
2. 論文標題 Investigation of enhanced Tacotron text-to-speech synthesis systems with self-attention for pitch accent language	5. 発行年 2019年
3. 雑誌名 2019 IEEE International Conference on Acoustics, Speech and Signal Processing	6. 最初と最後の頁 6905-6909
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/ICASSP.2019.8682353">https://doi.org/10.1109/ICASSP.2019.8682353</a>	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 -

1. 著者名 Lauri Juvola, Bajibabu Bollepalli, Junichi Yamagishi, Paavo Alku	4. 巻 -
2. 論文標題 Waveform generation for text-to-speech synthesis using pitch-synchronous multi-scale generative adversarial networks	5. 発行年 2019年
3. 雑誌名 2019 IEEE International Conference on Acoustics, Speech and Signal Processing	6. 最初と最後の頁 6915-6919
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/ICASSP.2019.8683271">https://doi.org/10.1109/ICASSP.2019.8683271</a>	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 該当する

1. 著者名 Fuming Fang, Xin Wang, Junichi Yamagishi, Isao Echizen	4. 巻 -
2. 論文標題 Audiovisual speaker conversion: jointly and simultaneously transforming facial expression and acoustic characteristics	5. 発行年 2019年
3. 雑誌名 2019 IEEE International Conference on Acoustics, Speech and Signal Processing	6. 最初と最後の頁 6795-6799
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/ICASSP.2019.8683872">https://doi.org/10.1109/ICASSP.2019.8683872</a>	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 -

1. 著者名 Shreyas Seshadri, Lauri Juvola, Junichi Yamagishi, Okko Rasanen, Paavo Alku	4. 巻 -
2. 論文標題 Cycle-consistent adversarial networks for non-parallel vocal effort based speaking style conversion	5. 発行年 2019年
3. 雑誌名 2019 IEEE International Conference on Acoustics, Speech and Signal Processing	6. 最初と最後の頁 6835-6839
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/ICASSP.2019.8682648">https://doi.org/10.1109/ICASSP.2019.8682648</a>	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 該当する

1. 著者名 Cassia Valentini-Botinhao, Junichi Yamagishi	4. 巻 8
2. 論文標題 Speech Enhancement of Noisy and Reverberant Speech for Text-to-Speech	5. 発行年 2018年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech, and Language Processing	6. 最初と最後の頁 1420-1433
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/TASLP.2018.2828980">https://doi.org/10.1109/TASLP.2018.2828980</a>	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 該当する

1. 著者名 Xin Wang, Shinji Takaki, Junichi Yamagishi	4. 巻 8
2. 論文標題 Autoregressive neural F0 model for statistical parametric speech synthesis	5. 発行年 2018年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech, and Language Processing	6. 最初と最後の頁 1406-1419
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/TASLP.2018.2828650">https://doi.org/10.1109/TASLP.2018.2828650</a>	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 -

1. 著者名 Tomi Kinnunen, Jaime Lorenzo-Trueba, Junichi Yamagishi, Tomoki Toda, Daisuke Saito, Fernando Villavicencio and Zhenhua Ling	4. 巻 -
2. 論文標題 A Spoofing Benchmark for the 2018 Voice Conversion Challenge: Leveraging from Spoofing Countermeasures for Speech Artifact Assessment	5. 発行年 2018年
3. 雑誌名 Speaker Odyssey 2018	6. 最初と最後の頁 187-194
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.21437/Odyssey.2018-27">http://dx.doi.org/10.21437/Odyssey.2018-27</a>	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 該当する

1. 著者名 Jaime Lorenzo-Trueba, Junichi Yamagishi, Tomoki Toda, Daisuke Saito, Fernando Villavicencio, Tomi Kinnunen and Zhenhua Ling	4. 巻 -
2. 論文標題 The Voice Conversion Challenge 2018: Promoting Development of Parallel and Nonparallel Methods	5. 発行年 2018年
3. 雑誌名 Speaker Odyssey 2018	6. 最初と最後の頁 195-202
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.21437/Odyssey.2018-28">http://dx.doi.org/10.21437/Odyssey.2018-28</a>	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 該当する

1. 著者名 Jaime Lorenzo-Trueba, Fuming Fang, Xin Wang, Isao Echizen, Junichi Yamagishi and Tomi Kinnunen	4. 巻 -
2. 論文標題 Can we steal your vocal identity from the Internet?: Initial investigation of cloning Obama's voice using GAN, WaveNet and low-quality found data	5. 発行年 2018年
3. 雑誌名 Speaker Odyssey 2018	6. 最初と最後の頁 240-247
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.21437/Odyssey.2018-34">http://dx.doi.org/10.21437/Odyssey.2018-34</a>	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 該当する

1. 著者名 Fuming Fang, Junichi Yamagishi, Isao Echizen, Jaime Lorenzo-Trueba	4. 巻 -
2. 論文標題 HIGH-QUALITY NONPARALLEL VOICE CONVERSION BASED ON CYCLE-CONSISTENT ADVERSARIAL NETWORK	5. 発行年 2018年
3. 雑誌名 2018 IEEE International Conference on Acoustics, Speech and Signal Processing	6. 最初と最後の頁 5279-5283
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/ICASSP.2018.8462342">https://doi.org/10.1109/ICASSP.2018.8462342</a>	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 -

1. 著者名 Gustav Eje Henter, Jaime Lorenzo-Trueba, Xin Wang, Mariko Kondo, Junichi Yamagishi	4. 巻 -
2. 論文標題 CYBORG SPEECH: DEEP MULTILINGUAL SPEECH SYNTHESIS FOR GENERATING SEGMENTAL FOREIGN ACCENT WITH NATURAL PROSODY	5. 発行年 2018年
3. 雑誌名 2018 IEEE International Conference on Acoustics, Speech and Signal Processing	6. 最初と最後の頁 4799-4803
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/ICASSP.2018.8462470">https://doi.org/10.1109/ICASSP.2018.8462470</a>	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 -



1. 著者名 Lauri Juvela, Bajibabu Bollepalli, Xin Wang, Hirokazu Kameoka, Manu Airaksinen, Junichi Yamagishi, Paavo Alku	4. 巻 -
2. 論文標題 SPEECH WAVEFORM SYNTHESIS FROM MFCC SEQUENCES WITH GENERATIVE ADVERSARIAL NETWORKS	5. 発行年 2018年
3. 雑誌名 2018 IEEE International Conference on Acoustics, Speech and Signal Processing	6. 最初と最後の頁 5679-5683
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/ICASSP.2018.8461852">https://doi.org/10.1109/ICASSP.2018.8461852</a>	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 該当する

1. 著者名 A COMPARISON OF RECENT WAVEFORM GENERATION AND ACOUSTIC MODELING METHODS FOR NEURAL-NETWORK-BASED SPEECH SYNTHESIS	4. 巻 -
2. 論文標題 A COMPARISON OF RECENT WAVEFORM GENERATION AND ACOUSTIC MODELING METHODS FOR NEURAL-NETWORK-BASED SPEECH SYNTHESIS	5. 発行年 2018年
3. 雑誌名 2018 IEEE International Conference on Acoustics, Speech and Signal Processing	6. 最初と最後の頁 4804-4808
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1109/ICASSP.2018.8461452">https://doi.org/10.1109/ICASSP.2018.8461452</a>	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 該当する

1. 著者名 Xin Wang, Shinji Takaki, Junichi Yamagishi	4. 巻 96
2. 論文標題 Investigating very deep highway networks for parametric speech synthesis	5. 発行年 2018年
3. 雑誌名 Speech Communication	6. 最初と最後の頁 1-9
掲載論文のDOI (デジタルオブジェクト識別子) <a href="https://doi.org/10.1016/j.specom.2017.11.002">https://doi.org/10.1016/j.specom.2017.11.002</a>	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 -

1. 著者名 Michael Pucher, Bettina Zillinger, Markus Toman, Junichi Yamagishi, Erich Schmid, Cassia Valentini-Botinhao, Dietmar Schabus, Thomas Woltron	4. 巻 46
2. 論文標題 Influence of speaker familiarity on blind and visually impaired children's perception of synthetic voices	5. 発行年 2017年
3. 雑誌名 Computer Speech & Language	6. 最初と最後の頁 179-195
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.1016/j.csl.2017.05.010">http://dx.doi.org/10.1016/j.csl.2017.05.010</a>	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 -
2. 論文標題 An RNN-based Quantized F0 Model with Multi-tier Feedback Links for Text-to-Speech Synthesis	5. 発行年 2017年
3. 雑誌名 Interspeech 2017	6. 最初と最後の頁 1059-1063
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.21437/Interspeech.2017-246">http://dx.doi.org/10.21437/Interspeech.2017-246</a>	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 -

1. 著者名 Shinji Takaki, Hirokazu Kameoka, Junichi Yamagishi	4. 巻 -
2. 論文標題 Direct modeling of frequency spectra and waveform generation based on phase recovery for DNN-based speech synthesis	5. 発行年 2017年
3. 雑誌名 Interspeech 2017	6. 最初と最後の頁 1128-1132
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.21437/Interspeech.2017-488">http://dx.doi.org/10.21437/Interspeech.2017-488</a>	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 -

1. 著者名 Toru Nakashika, Shinji Takaki, Junichi Yamagishi	4. 巻 -
2. 論文標題 Complex-valued restricted Boltzmann machine for direct learning of frequency spectra	5. 発行年 2017年
3. 雑誌名 Interspeech 2017	6. 最初と最後の頁 4021-4025
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.21437/Interspeech.2017-584">http://dx.doi.org/10.21437/Interspeech.2017-584</a>	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 -

1. 著者名 Lauri Juvola, Bajibabu Bollepalli, Junichi Yamagishi, Paavo Alku	4. 巻 -
2. 論文標題 Reducing mismatch in training of DNN-based glottal excitation models in a statistical parametric text-to-speech system	5. 発行年 2017年
3. 雑誌名 Interspeech 2017	6. 最初と最後の頁 1368-1372
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.21437/Interspeech.2017-848">http://dx.doi.org/10.21437/Interspeech.2017-848</a>	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 該当する

1. 著者名 Takuhiro Kaneko, Shinji Takaki, Hirokazu Kameoka, Junichi Yamagishi	4. 巻 -
2. 論文標題 Generative Adversarial Network-based Postfilter for STFT Spectrograms	5. 発行年 2017年
3. 雑誌名 Interspeech 2017	6. 最初と最後の頁 3389-3393
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.21437/Interspeech.2017-962">http://dx.doi.org/10.21437/Interspeech.2017-962</a>	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 -

1. 著者名 M. Sam Ribeiro, Oliver Watts, Junichi Yamagishi	4. 巻 -
2. 論文標題 Learning word vector representations based on acoustic counts	5. 発行年 2017年
3. 雑誌名 Interspeech 2017	6. 最初と最後の頁 799-803
掲載論文のDOI (デジタルオブジェクト識別子) <a href="http://dx.doi.org/10.21437/Interspeech.2017-1340">http://dx.doi.org/10.21437/Interspeech.2017-1340</a>	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 該当する

[学会発表] 計22件 (うち招待講演 1件 / うち国際学会 10件)

1. 発表者名 Igor Jauk, Jaime Lorenzo-Trueba, Junichi Yamagishi, Antonio Bonafonte
2. 発表標題 Expressive Speech Synthesis Using Sentiment Embeddings
3. 学会等名 Interspeech 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Lauri Juvela, Vassilis Tsirias, Bajibabu Bollepalli, Manu Airaksinen, Junichi Yamagishi, Paavo Alku
2. 発表標題 Speaker-independent Raw Waveform Model for Glottal Excitation
3. 学会等名 Interspeech 2018 (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 Multimodal Speech Synthesis Architecture for Unsupervised Speaker Adaptation
3. 学会等名 Interspeech 2018 ( 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 Spoken Language Technology Workshop (SLT) ( 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 ( 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 ( 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 ( 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 ( 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 ( 国際学会 )
4 . 発表年 2019年

1 . 発表者名 Shreyas Seshadri, Lauri Juvola, Junichi Yamagishi, Okko Rasanen, Paavo Alku
2 . 発表標題 Cycle-consistent adversarial networks for non-parallel vocal effort based speaking style conversion
3 . 学会等名 2019 IEEE International Conference on Acoustics, Speech and Signal Processing ( 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 Investigation of WaveNet for Text-to-Speech Synthesis
3. 学会等名 情報処理学会 第120回 音声言語情報処理研究会
4. 発表年 2018年

1. 発表者名 Jaime Lorenzo-Trueba, Xin Wang, Junichi Yamagishi
2. 発表標題 Stealing your vocal identity from the internet: cloning Obama's voice from found data using GAN and Wavenet
3. 学会等名 情報処理学会 第120回 音声言語情報処理研究会
4. 発表年 2018年

1. 発表者名 Gustav Henter, Jaime Lorenzo-Trueba, Xin Wang, Kondo Mariko, Junichi Yamagishi
2. 発表標題 Generating segment-level foreign-accented synthetic speech with natural speech prosody
3. 学会等名 情報処理学会 第120回 音声言語情報処理研究会
4. 発表年 2018年

1. 発表者名 中鹿 亘, 高木 信二, 山岸 順一
2. 発表標題 リカレント構造を持つ複素制限ボルツマンマシンによる複素スペクトル系列モデリング
3. 学会等名 情報処理学会 第120回 音声言語情報処理研究会
4. 発表年 2018年

1. 発表者名 房 福明, Jaime Lorenzo- Trueba, 山岸 順一, 越前 功
2. 発表標題 CycleGANを用いたクロスリンガル声質変換
3. 学会等名 情報処理学会 第120回 音声言語情報処理研究会
4. 発表年 2018年

1. 発表者名 房 福明, 山岸 順一, 越前 功
2. 発表標題 CycleGANを用いた高品質なノンパラレル声質変換
3. 学会等名 第19回音声言語シンポジウム
4. 発表年 2017年

1. 発表者名 Jaime Lorenzo-Trueba · Gustav Henter · Shinji Takaki · Junichi Yamagishi
2. 発表標題 Analyzing the impact of including listener perception annotations in RNN-based emotional speech synthesis
3. 学会等名 第19回音声言語シンポジウム
4. 発表年 2017年

1. 発表者名 山岸順一
2. 発表標題 ASVspoof: 話者照合における生体検知
3. 学会等名 第7回バイオメトリクスと認識・認証シンポジウム(招待講演)
4. 発表年 2017年

1. 発表者名 中鹿亘, 高木信二, 山岸順一
2. 発表標題 複素RBMを用いた音声スペクトルモデリングの改良と評価
3. 学会等名 日本音響学会2017年秋季研究発表会
4. 発表年 2017年

1. 発表者名 Xin Wang, Shinji Takaki, Junichi Yamagishi
2. 発表標題 Autoregressive quantized F0 modeling using a recurrent neural network with feedback links
3. 学会等名 2017年8月音声研究会
4. 発表年 2017年

1. 発表者名 中鹿亘, 高木信二, 山岸順一
2. 発表標題 複素RBM: 制限ボルツマンマシンの複素数拡張と音声信号への応用と評価
3. 学会等名 第117回音声言語情報処理研究会
4. 発表年 2017年

1. 発表者名 金子卓弘, 高木信二, 亀岡弘和, 山岸順一
2. 発表標題 敵対的学習に基づくSTFTスペクトログラムのポストフィルタリング
3. 学会等名 2017年6月度音声研究会
4. 発表年 2017年



〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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
--	---------------------------	-----------------------	----