

令和 3 年 5 月 31 日現在

機関番号：14501

研究種目：基盤研究(B)（一般）

研究期間：2017～2020

課題番号：17H01995

研究課題名（和文）構音障がい者の複数モダリティを用いたコミュニケーション支援技術の研究

研究課題名（英文）Research on communication support technology for persons with articulation disorders using multiple modalities

研究代表者

滝口 哲也（TAKIGUCHI, Tetsuya）

神戸大学・都市安全研究センター・教授

研究者番号：40397815

交付決定額（研究期間全体）：（直接経費） 13,200,000円

研究成果の概要（和文）：本研究では構音障がい者のコミュニケーション支援技術として、3つの項目について研究を遂行した。(1)構音障がい者の音声認識手法として、複数障がい者音声データベースを活用した音声認識モデル学習方法などを提案し有効性を示した。(2)音声生成手法として、健常者音声合成モデルを活用した話者性を維持した聞き取り容易な音声生成システムなどを提案し有効性を示した。(3)マルチモーダル発話認識手法として、Lip-readingモデルの教師なしドメイン適応などについて提案し有効性を示した。

研究成果の学術的意義や社会的意義

近年、スマートスピーカーなど音声を用いた端末入力サービスが発表されている。しかしこれらのサービスでは、明瞭な発話ができる人を対象としており、発話困難者には対応していない。本研究では、構音障がい者の音声コミュニケーション支援技術として、複数データベース（複数モデル）を活用した音声認識、音声合成を提案した。得られた発話コミュニケーション支援技術に関する成果は、障がい者の自立生活支援などに資するものである。

研究成果の概要（英文）：In this study, we focused on three subjects with respect to communication assistive technologies for persons with articulation disorders. (1) For automatic speech recognition, we proposed an end-to-end system using multiple speech databases. (2) For speech synthesis, we proposed a hybrid approach, using two models of a physically unimpaired person and a person with an articulation disorder to generate an intelligible voice while preserving the speaker's individuality. (3) For multi-modal speech recognition, we proposed unsupervised domain adaptation to lip reading.

研究分野：メディア情報工学

キーワード：ヒューマン・インターフェース

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

人は社会において様々な人とのコミュニケーションや交わりを通して成長し、生きがいを見出していく。しかしながら障がい者においては、社会に出たくても例えば運動機能の劣化のため自分自身では屋外に出るのが難しい場合がある。また構音劣化のため発話も不明瞭になり、自分の意思を声にして伝えることが困難であり、さらに身体的な障がいも伴っている場合は、携帯電話等の端末操作も困難になり、コミュニケーション社会から遠ざからざるを得ない状況にある。そのような人々を取り残すような社会ではなく、情報技術を活用して、情報弱者の生活をサポートし、温かく迎え入れることのできる「全ての人が共生したコミュニケーション社会」「生きがいのある豊かなコミュニケーション社会」の構築が、社会基盤システムとして必要である。

### 2. 研究の目的

構音障がいの人々は一人一人その障がい特性が多様であり、コミュニケーションが困難な場合がある。そのため一つのモダリティだけではなく、複数のモダリティ(音、映像)に基づくユニバーサルコミュニケーション支援技術が必須である。本研究では発話コミュニケーション障がい者独自の課題に対して、実証実験まで行う。具体的には、聞き取り困難な発話に対して、(1) 深層学習を軸にした発話意図理解、(2) 障がい者の話者性を維持した音声生成、(3) 顔部位などの特徴を利用したマルチモーダル認識を実現する新たなアルゴリズムを提案し、障がい者の身体能力の壁を超越することができるコミュニケーション支援技術の実現を目指す。

### 3. 研究の方法

深層学習に基づく音声、画像処理技術を基盤として、構音障がい者のユニバーサルコミュニケーションを促進するための支援技術の研究開発を行う。多様な障がい者の課題に対応するため、3つのサブテーマを設定し研究を遂行する。

#### (1) 障がい者の発話意図理解

複数障がい者音声データベースを活用した新たな End-to-End 音声認識モデルの学習方法、健常者の不特定話者音響モデルを活用した障がい者音響モデルへの適応手法などについて研究する。

#### (2) 障がい者の話者性を維持した音声生成

モデル学習時にパラレルデータ(入力話者と出力話者の同一テキスト発話)を使用しない Exemplar-based 声質変換手法、健常者音声合成モデルと障がい者音声合成モデルの統合による話者性を維持した聞き取り容易な音声生成システムなどについて研究する。

#### (3) 顔部位などの特徴量を考慮したマルチモーダル発話処理

Lip-reading モデルの教師なしドメイン適応、雑音環境下における音声画像統合ニューラルネットワークなどについて研究する。

### 4. 研究成果

本研究で得られた主な研究成果は以下のとおりである。

#### (1) 複数障がい者音声データベースを活用した音声認識

本研究では、評価対象である日本人構音障がい者の音声だけでなく、外国人構音障がい者と日本人健常者の音声をを用いたデータ拡張を提案した。障がいの有無に関わらず、言語が同じであれば言語モデルは共有できると考えられる。そこで提案手法では、構音障がいを持つ日本人話者と外国人話者で音響ネットワークを共有し、日本人障がい者と日本人健常者で言語ネットワークを共有する構造を持つネットワークを構築した(図1)。

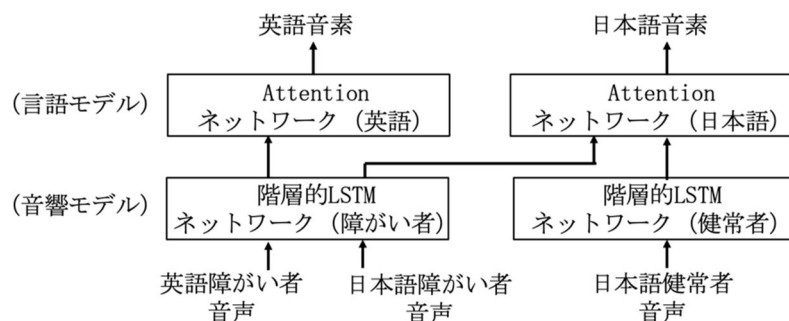


図1. 複数データベースを用いた End-to-End 音声認識ネットワーク

図 2 に日本語構音障がい者の音声認識結果（文字誤り率）を示す．End-to-End 音声認識ニューラルネットワークの学習データとして，認識対象の日本語構音障がい者のみならず，日本語健常者発話データ，さらに英語構音障がい者の発話データも用いることで認識精度が改善された．

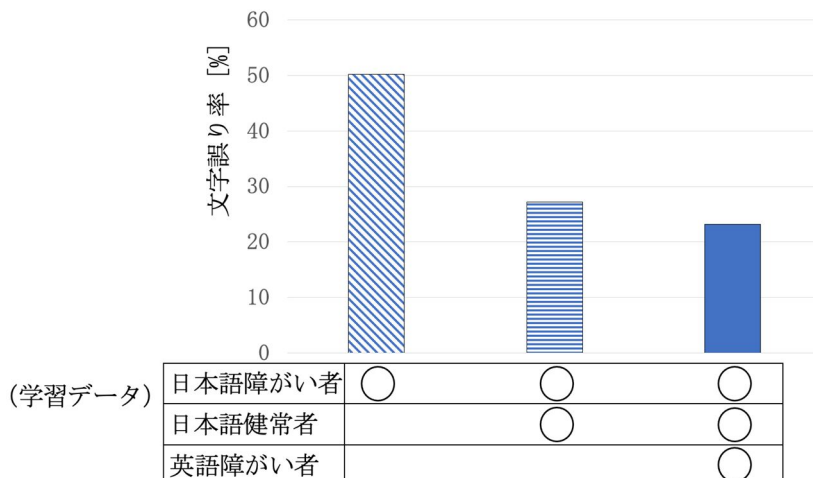


図 2. 障がい者の音声認識結果

## (2) 雑談対話システム

発話意図支援システムのツールとして，本研究では新たに雑談対話システムの構築を行った．従来の雑談システムでは，浅い知識しか組み込まれていなかったが，本システムではユーザ発話を掘り下げるような，深い知識を含んだ応答生成を行うモデルをマルチタスク学習により実現し有効性を示した．

## (3) 顔画像特徴量を用いた音声変換

音声情報が含まれていない唇動画像からその唇の動きに対応した音声へと変換する音声生成システムの構築を行った．従来手法の GMM(Gaussian Mixture Model)を用いた唇画像から音声への変換では，過剰な平滑化の問題，さらに線形変換手法であるため人間の声の特徴を表すスペクトルの推定精度に問題があった．それらの課題を解決するため，Bi-LSTM(Bidirectional Long Short Term Memory)を用いた推定手法を提案し，「聞き取りやすさ」調査において提案手法の有効性を示した．

## (4) 音声合成における言語特徴量の検討

脳性麻痺者を対象として，音声合成を用いて彼らのコミュニケーションを支援するシステムを提案した．脳性麻痺者の収録音声には，読み上げ時の筋肉の不随意運動により，同一のプレズグループやアクセント句の中に多くの息詰まりなどが存在する．よって，テキストを形態素解析して得られる言語特徴量と学習に用いる音声との間にミスマッチが存在する．本研究では学習及び推定に用いる言語特徴量に使用する特徴を検討することで，より明瞭性の高い音声を作成するシステムを提案した．主観評価実験により話者性と明瞭度の試験を行った結果，提案したモデルで作成した音声は，話者性を維持しており，録音音声と比較して高い明瞭度を示した．

## (5) Exemplar-based 手法に基づく声質変換手法

従来はモデル学習時にパラレルデータを必要としていた．パラレルデータとは，入力話者と出力話者の同一発話内容による音声対であり，パラレルデータの作成には様々な制限が課せられる．第一に，発話データは同一の発話内容でないといけないという制限があるため，選択（または作成）できる学習データセットの自由度は低い．第二に，フレーム単位で入出力音声の同期を取る必要があるため，動的計画法などを用いてアライメントを取るが，完全にフレームの同期が取れている保証がない．したがって伸縮の際に音声に変換が加わっているなどの問題があった．Exemplar-based 手法において，変換辞書はパラレルデータから構成されるため，パラレルデータのアライメント誤差が声質変換性能に悪影響を及ぼす可能性がある．本研究ではパラレルデータを使用しない非負値タッカー分解に基づく辞書学習法を提案し有効性を

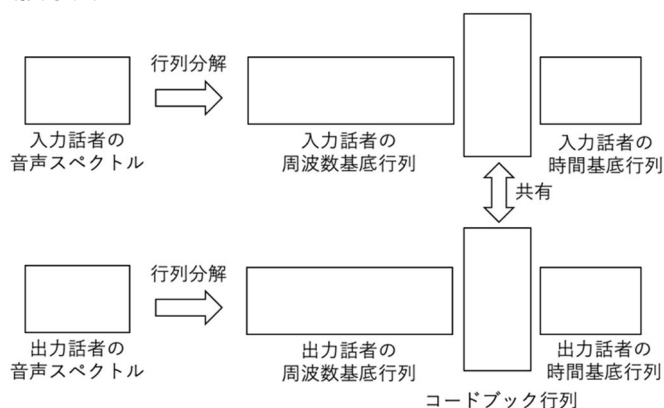


図 3. 非負値タッカー分解によるパラレル辞書学習

を提案し有効性を示した．

示した。

#### (6) 音源復元のための物体振動抽出手法の検討

音響信号の影響により発生した物体の微細振動をハイスピードカメラで捉えることにより、映像から音響信号を復元する手法を検討した。本研究では振幅の大きな画像領域、位相差の大きい画像領域に注目した振動抽出手法、および畳み込みニューラルネットワークに基づく音変換手法を提案し、実験によりその効果を確認した。

#### (7) 知識蒸留に基づく Lip reading のための教師なしドメイン適応

ドメイン適応の目的は、元ドメインで学習されたモデルを、できる限り少量の適応データを用いて新しい目標ドメインに適応することである。特に、適応データのラベル情報を全く使用できない場合は、教師なしドメイン適応 (unsupervised domain adaptation: UDA) と呼ばれる。従来のドメイン適応手法では適応データと学習データが同じクラス集合に属していることを前提としているため、未知クラスに属するデータを適応データとして使用することができない。より実用的なドメイン適応のために、このような未知クラスのデータを利用できる適応手法が求められる。そこで本研究では、未知クラスのデータを用いたドメイン適応手法として、lip reading のためのクロスモーダルな知識蒸留に基づくアプローチを提案した。具体的には学習済みの音声認識モデル(教師モデル)の振る舞いを模倣するように対象 lip reading モデル(生徒モデル)を学習する。一般に発話認識において、音声は唇の動きよりも、より多くの情報を持っているとされ、音声認識の精度は lip reading よりも高くなる。そのため、音声認識モデルの出力は強力な教師信号となると期待される。単語認識実験により、提案手法が教師なしドメイン適応の性能を向上させ、既知クラスを用いたドメイン適応と同等の性能を持つことが示された。

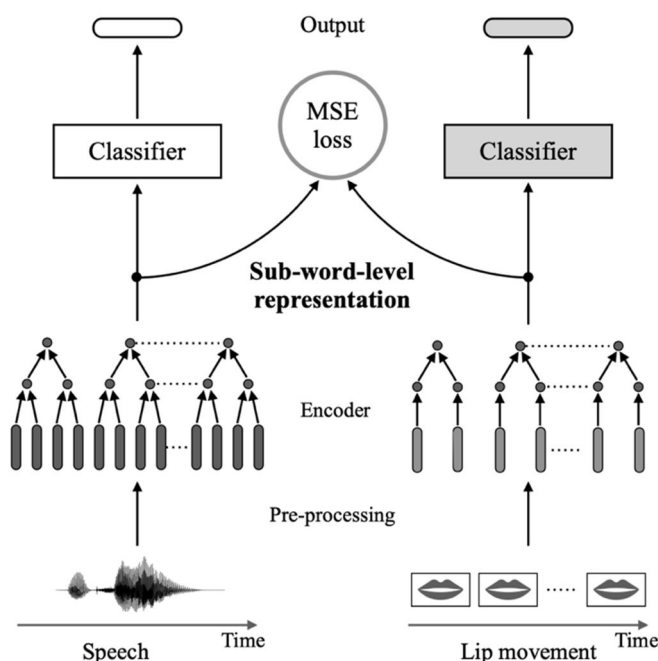


図 4. 音声モデルと唇画像モデルのクロスモーダル知識統合

#### (8) 局所アテンションを用いた音声画像統合

近年、深層学習を用いた音声認識システムの精度が大幅に向上している。しかし、音声信号が劣化するような雑音環境下では、発話内容の認識精度が低下するという問題がある。本研究では音声と口唇動画の統合による発話認識の精度向上を行った。従来手法の一つに音響特徴量をクエリとして全ての時刻の画像特徴量にアテンションをかけるグローバルアテンションがある。一般にアテンションは雑音の大きい観測音声では重みの推定が困難であることが知られているため、グローバルアテンションでは雑音環境下において正確な重みの推定が困難な可能性がある。そこで本研究では、ある時刻の音声に関する画像のフレームは全体の一部であるという考えに基づき、重みの計算を一部の画像フレームに限定する局所アテンションを提案した。局所アテンションであれば雑音環境下においても重み計算を行うフレームを限定することで、適切な重みの推定が可能であることが期待される。

評価実験の結果、提案手法は従来手法と比較して、背景雑音が存在する環境下で従来手法よりも優れた結果を示した。また、妨害発話が存在する難易度の高いタスクにおいても、提案手法が従来のグローバルアテンションを上回る認識精度を示した。

## 5. 主な発表論文等

〔雑誌論文〕 計15件（うち査読付論文 15件 / うち国際共著 1件 / うちオープンアクセス 6件）

1. 著者名 Yuki Takashima, Ryoichi Takashima, Tetsuya Takiguchi, Yasuo Ariki	4. 巻 7
2. 論文標題 Knowledge Transferability Between the Speech Data of Persons With Dysarthria Speaking Different Languages for Dysarthric Speech Recognition	5. 発行年 2019年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 164320-164326
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ACCESS.2019.2951856	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 Yuki Takashima, Toru Nakashika, Tetsuya Takiguchi, and Yasuo Ariki	4. 巻 17
2. 論文標題 Non-parallel dictionary learning for voice conversion using non-negative Tucker decomposition	5. 発行年 2019年
3. 雑誌名 EURASIP Journal on Audio, Speech, and Music Processing	6. 最初と最後の頁 11 pages
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s13636-019-0160-1	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 Zhaojie Luo, Jinhui Chen, Tetsuya Takiguchi, Yasuo Ariki	4. 巻 27
2. 論文標題 Emotional Voice Conversion Using Dual Supervised Adversarial Networks With Continuous Wavelet Transform F0 Features	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 1535-1548
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TASLP.2019.2923951	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 Zhaojie Luo, Jinhui Chen, Tetsuya Takiguchi, and Yasuo Ariki	4. 巻 8
2. 論文標題 Neutral-to-emotional voice conversion with cross-wavelet transform F0 using generative adversarial networks	5. 発行年 2019年
3. 雑誌名 APSIPA Transactions on Signal and Information Processing	6. 最初と最後の頁 1-11
掲載論文のDOI (デジタルオブジェクト識別子) 10.1017/ATSIP.2019.3	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 Tristan Hascoet, Yasuo Arika and Tetsuya Takiguchi	4. 巻 2019: 13
2. 論文標題 Semantic embeddings of generic objects for zero-shot learning	5. 発行年 2019年
3. 雑誌名 EURASIP Journal on Image and Video Processing	6. 最初と最後の頁 1-14
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s13640-018-0371-x	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 Jinhui Chen, Zhaojie Luo, Zhihong Zhang, Faliang Huang, Zhiling Ye, Tetsuya Takiguchi, Edwin R. Hancock	4. 巻 21
2. 論文標題 Polar Transformation on Image Features for Orientation-Invariant Representations	5. 発行年 2018年
3. 雑誌名 IEEE Transactions on Multimedia	6. 最初と最後の頁 300-313
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TMM.2018.2856121	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 Yuki Takashima, Tetsuya Takiguchi, Yasuo Arika	4. 巻
2. 論文標題 Exemplar-based Lip-to-Speech Synthesis Using Convolutional Neural Networks	5. 発行年 2019年
3. 雑誌名 International Workshop on Frontiers of Computer Vision	6. 最初と最後の頁 5 pages
掲載論文のDOI (デジタルオブジェクト識別子) なし	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 Yohei Fuse, Yusuke Yasumi, Tetsuya Takiguchi	4. 巻
2. 論文標題 Sound Recovery Considering the Vibration Direction of an Object in a Video	5. 発行年 2018年
3. 雑誌名 IEEE ISM	6. 最初と最後の頁 171-174
掲載論文のDOI (デジタルオブジェクト識別子) なし	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 Yuki Matsuyoshi, Tetsuya Takiguchi, Yasuo Arika	4. 巻
2. 論文標題 User's Intention Understanding in Question-Answering System Using Attention-based LSTM	5. 発行年 2018年
3. 雑誌名 APSIPA	6. 最初と最後の頁 1752-1755
掲載論文のDOI (デジタルオブジェクト識別子) なし	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 Ryuka Nanzaka, Tetsuya Takiguchi	4. 巻
2. 論文標題 Hybrid Text-to-Speech for Articulation Disorders with a Small Amount of Non-Parallel Data	5. 発行年 2018年
3. 雑誌名 APSIPA	6. 最初と最後の頁 1761-1765
掲載論文のDOI (デジタルオブジェクト識別子) なし	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 Kazuaki Furumai, Tetsuya Takiguchi, Yasuo Arika	4. 巻
2. 論文標題 Chat Response Generation Based on Semantic Prediction Using Distributed Representations of Words	5. 発行年 2018年
3. 雑誌名 International Workshop on Spoken Dialog System Technology	6. 最初と最後の頁 6 pages
掲載論文のDOI (デジタルオブジェクト識別子) なし	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 Rikito Marumoto, Katsuyuki Tanaka, Tetsuya Takiguchi, Yasuo Arika	4. 巻
2. 論文標題 Debate Dialog for News Question Answering System 'NetTv' -Debate Based on Claim and Reason Estimation-	5. 発行年 2018年
3. 雑誌名 International Workshop on Spoken Dialog System Technology	6. 最初と最後の頁 8 pages
掲載論文のDOI (デジタルオブジェクト識別子) なし	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 Yuki Takashima, Hajime Yano, Toru Nakashika, Tetsuya Takiguchi, Yasuo Arika	4. 巻
2. 論文標題 Parallel-Data-Free Dictionary Learning for Voice Conversion Using Non-Negative Tucker Decomposition	5. 発行年 2018年
3. 雑誌名 ICASSP	6. 最初と最後の頁 5294-5298
掲載論文のDOI (デジタルオブジェクト識別子) なし	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 Jinhui Chen, Tetsuya Takiguchi, Yasuo Arika	4. 巻 11
2. 論文標題 Rotation-reversal Invariant HOG Cascade for Facial Expression Recognition	5. 発行年 2017年
3. 雑誌名 Signal, Image and Video Processing	6. 最初と最後の頁 1485-1492
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s11760-017-1111-x	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 Luo Zhaojie, Chen Jinhui, Takiguchi Tetsuya, Arika Yasuo	4. 巻 2017
2. 論文標題 Emotional voice conversion using neural networks with arbitrary scales F0 based on wavelet transform	5. 発行年 2017年
3. 雑誌名 EURASIP Journal on Audio, Speech, and Music Processing	6. 最初と最後の頁 1-13
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s13636-017-0116-2	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50件 (うち招待講演 2件 / うち国際学会 9件)

1. 発表者名 古舞 千咲, 有木 康雄, 滝口 哲也
2. 発表標題 口バストな対話システムのための発話スタイルに非依存なエンコード・デコード手法の提案
3. 学会等名 人工知能学会 言語・音声理解と対話処理研究会
4. 発表年 2019年



1. 発表者名 高島 悠樹, 高島 遼一, 滝口 哲也, 有木 康雄
2. 発表標題 構音障害者音声認識のための健常者音声及び他言語障害者音声を用いた転移学習
3. 学会等名 電子情報通信学会技術研究報告 音声研究会
4. 発表年 2019年

1. 発表者名 麻生 大聖, 高島 遼一, 滝口 哲也, 有木 康雄
2. 発表標題 WordNetを用いた雑談対話システムの汎化性能の向上
3. 学会等名 電子情報通信学会技術研究報告
4. 発表年 2019年

1. 発表者名 片平 健太, 足立 優司, 田井 清登, 高島 遼一, 滝口 哲也
2. 発表標題 Bidirectional Gated Recurrent Units を用いた歌声合成に関する検討
3. 学会等名 情報処理学会 音声言語情報処理研究会
4. 発表年 2019年

1. 発表者名 古舞千暁, 滝口哲也, 有木康雄
2. 発表標題 発話スタイルに非依存な文章表現を用いたスタイルロバストなSeq2seqモデルの提案
3. 学会等名 言語処理学会第26回年次大会発表論文集
4. 発表年 2020年

1. 発表者名 長谷川公大, 松本剛明, 高島遼一, 滝口哲也, 有木康雄, 三田村照子
2. 発表標題 Transfer Learning to Generate Multiple Sentence Question with Leveraging Difference between Datasets
3. 学会等名 言語処理学会第26回年次大会発表論文集
4. 発表年 2020年

1. 発表者名 薛強, 滝口哲也, 有木康雄
2. 発表標題 感情・欲求に基づく主観性を持つ雑談対話システム
3. 学会等名 言語処理学会第26回年次大会発表論文集
4. 発表年 2020年

1. 発表者名 松原圭亮, 岡本拓磨, 高島遼一, 滝口哲也, 戸田智基, 志賀芳則, 河井 恒
2. 発表標題 リアルタイムニューラルボコーダにおける学習データ量の影響の調査
3. 学会等名 日本音響学会2020年春季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 高島 悠樹, 相原 龍, 高島 遼一, 滝口 哲也, 有木 康雄, 村山 修
2. 発表標題 クロスモーダル知識蒸留に基づくLip readingのための教師なしドメイン適応
3. 学会等名 日本音響学会2020年春季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 澤 佑哉, 高島 遼一, 滝口 哲也, 有木 康雄
2. 発表標題 Hybrid CTC/attentionモデルを用いた構音障害者音声認識の検討
3. 学会等名 日本音響学会2020年春季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 南坂 竜翔, 高島 遼一, 滝口 哲也
2. 発表標題 少量データを用いた構音障害者音声合成の健常者モデルによる明瞭性改善
3. 学会等名 日本音響学会2020年春季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 片平 健太, 足立 優司, 田井 清登, 高島 遼一, 滝口 哲也
2. 発表標題 自由な歌唱速度の歌声の合成に関する検討
3. 学会等名 日本音響学会2020年春季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 黄 伊莎, Tristan Hascoet, 高島遼一, 滝口哲也, 有木康雄
2. 発表標題 Differentiable Programmingを用いた強化学習の最適化
3. 学会等名 情報処理学会第82回全国大会講演論文集
4. 発表年 2020年

1. 発表者名 長谷川貴大, Tristan Hascoet, 高島遼一, 滝口哲也, 有木康雄
2. 発表標題 ニューロンセグメンテーションにおけるマルチドメイン学習による汎化性能の改善
3. 学会等名 情報処理学会第82回全国大会講演論文集
4. 発表年 2020年

1. 発表者名 麻生 大聖, 高島 遼一, 滝口 哲也, 有木 康雄
2. 発表標題 外部知識を用いた雑談対話システムの汎化性能向上の検討
3. 学会等名 日本音響学会2019年秋季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 高島 悠樹, 相原 龍, 高島 遼一, 滝口 哲也, 有木 康雄, 村山 修
2. 発表標題 Lip readingのためのクロスモーダルTeacher-Student学習
3. 学会等名 日本音響学会2019年秋季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 Zhaojie Luo, Ryoichi Takashima, Tetsuya Takiguchi, Yasuo Arika
2. 発表標題 Speech-to-Speech Translation using Dual Learning and Prosody Conversion
3. 学会等名 日本音響学会2019年秋季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 南坂 竜翔, 高島 遼一, 滝口 哲也, 有木 康雄
2. 発表標題 構音障害者の少量データを用いた深層学習による音声合成の検討
3. 学会等名 日本音響学会2019年秋季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 片平 健太, 足立 優司, 田井 清登, 高島 遼一, 滝口 哲也
2. 発表標題 歌声の母音変化を考慮した歌声合成の検討
3. 学会等名 日本音響学会2019年秋季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 高島 遼一, 滝口 哲也, 有木 康雄
2. 発表標題 構音障害者を対象とした日本語大語彙連続音声認識の検討
3. 学会等名 日本音響学会2019年秋季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 Tristan E. M. Hascoet, Yasuo Arika, Tetsuya Takiguchi
2. 発表標題 On zero-shot recognition of generic objects
3. 学会等名 第22回画像の認識・理解シンポジウム(招待講演)
4. 発表年 2019年

1. 発表者名 Weihao Zhuang, Tristan Hascoet, Ryoichi Takashima, Tetsuya Takiguchi, Yasuo Arika
2. 発表標題 Reduce GPU Memory Usage of Training Neural Network by CPU Offloading
3. 学会等名 第22回画像の認識・理解シンポジウム
4. 発表年 2019年

1. 発表者名 滝口 哲也
2. 発表標題 発話障害者の音声コミュニケーション支援
3. 学会等名 電子情報通信学会技術研究報告（招待講演）
4. 発表年 2018年

1. 発表者名 北村 毅, 滝口 哲也
2. 発表標題 音響特徴量補正による構音障害者を対象としたDNN音声合成
3. 学会等名 日本音響学会2019年春季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 Zhaojie Luo, Tetsuya Takiguchi, Yasuo Arika
2. 発表標題 Speech Prosody Conversion using Sequence Generative Adversarial Nets with Continuous Wavelet Transform F0 features
3. 学会等名 日本音響学会2019年春季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 麻生 大聖, 滝口 哲也, 有木 康雄
2. 発表標題 マルチタスク学習による雑談対話システムへの知識付与
3. 学会等名 日本音響学会2019年春季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 松好 祐紀, 滝口 哲也, 有木 康雄
2. 発表標題 ユーザーの発話意図理解に基づくインタビュー発話の生成
3. 学会等名 日本音響学会2019年春季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 谷田 啓一, Tristan Hascoet, 滝口 哲也, 有木 康雄
2. 発表標題 ゼロショット学習を用いた一般物体セグメンテーション
3. 学会等名 情報処理学会第81回全国大会講演論文集
4. 発表年 2019年

1. 発表者名 南坂 竜翔, 滝口 哲也, 有木 康雄
2. 発表標題 CycleGANに基づくノンパラレル声質変換を用いた構音障害者音声合成
3. 学会等名 日本音響学会2018年秋季研究発表会講演論文集
4. 発表年 2018年

1. 発表者名 羅 兆傑, 滝口 哲也, 有木 康雄
2. 発表標題 Neutral-to-Emotional Voice Conversion with Latent Representations of F0 using Generative Adversarial Networks
3. 学会等名 日本音響学会2018年秋季研究発表会講演論文集
4. 発表年 2018年

1. 発表者名 高島 悠樹, 滝口 哲也, 有木 康雄
2. 発表標題 非負値行列因子分解に基づく構音障害者音声の高域付加の検討
3. 学会等名 日本音響学会2018年秋季研究発表会講演論文集
4. 発表年 2018年

1. 発表者名 高島悠樹, 滝口 哲也, 有木 康雄
2. 発表標題 ハイスピードカメラ画像を用いた唇動画像からの音声生成
3. 学会等名 第21回画像の認識・理解シンポジウム
4. 発表年 2018年

1. 発表者名 布施 陽平, 滝口 哲也, 有木 康雄
2. 発表標題 物体振動を用いた畳み込みニューラルネットワークによる音源復元
3. 学会等名 第21回画像の認識・理解シンポジウム
4. 発表年 2018年



1. 発表者名 松好祐紀, 滝口 哲也, 有木康雄
2. 発表標題 Attention-based LSTMを用いた意図理解とキーワード抽出の統合による質問応答システム
3. 学会等名 電子情報通信学会技術研究報告
4. 発表年 2018年

1. 発表者名 Rina Ra, Ryo Aihara, Tetsuya Takiguchi, Yasuo Arika
2. 発表標題 Visual-to-Speech Conversion Based on Maximum Likelihood Estimation
3. 学会等名 IAPR International Conference on Machine Vision Applications (国際学会)
4. 発表年 2017年

1. 発表者名 Zhaojie Luo, Jinhui Chen, Tetsuya Takiguchi, Yasuo Arika
2. 発表標題 FACIAL EXPRESSION RECOGNITION WITH DEEP AGE
3. 学会等名 The Second Workshop on Human Identification in Multimedia (国際学会)
4. 発表年 2017年

1. 発表者名 Tsuyoshi Kitamura, Tetsuya Takiguchi, Yasuo Arika, Kiyohiro Omori
2. 発表標題 Individuality-Preserving Speech Synthesis System for Hearing Loss Using Deep Neural Networks
3. 学会等名 1st International Workshop on Challenges in Hearing Assistive Technology (国際学会)
4. 発表年 2017年

1. 発表者名 Zhaojie Luo, Jinhui Chen, Tetsuya Takiguchi, Yasuo Ariki
2. 発表標題 Emotional Voice Conversion with Adaptive Scales F0 based on Wavelet Transform using Limited Amount of Emotional Data
3. 学会等名 Interspeech (国際学会)
4. 発表年 2017年

1. 発表者名 李 権俊, 滝口 哲也, 有木 康雄
2. 発表標題 深層学習による位相情報を考慮した音声合成の検討
3. 学会等名 日本音響学会2017年秋季研究発表会講演論文集
4. 発表年 2017年

1. 発表者名 北村 毅, 滝口 哲也, 有木 康雄
2. 発表標題 深層学習を用いた構音障害者の音声合成
3. 学会等名 日本音響学会2017年秋季研究発表会講演論文集
4. 発表年 2017年

1. 発表者名 羅 兆傑, 滝口 哲也, 有木 康雄
2. 発表標題 EMOTIONAL VOICE CONVERSION WITH WAVELET TRANSFORM USING DUAL SUPERVISED ADVERSARIAL NETWORKS
3. 学会等名 日本音響学会2018年春季研究発表会講演論文集
4. 発表年 2018年

1. 発表者名 羅 里奈, 滝口 哲也, 有木 康雄
2. 発表標題 顔画像特徴量を用いた統計的手法によるF0推定
3. 学会等名 日本音響学会2018年春季研究発表会講演論文集
4. 発表年 2018年

1. 発表者名 南坂 竜翔, 滝口 哲也, 有木 康雄
2. 発表標題 構音障害者の少量学習データによる音声合成の検討
3. 学会等名 日本音響学会2018年春季研究発表会講演論文集
4. 発表年 2018年

1. 発表者名 北村 毅, 滝口 哲也, 有木 康雄
2. 発表標題 構音障害者を対象としたDNN音声合成に関する言語特徴量の検討
3. 学会等名 日本音響学会2018年春季研究発表会講演論文集
4. 発表年 2018年

1. 発表者名 安見 祐亮, 滝口 哲也, 有木 康雄
2. 発表標題 ハイスピード映像からの音源復元のための物体振動抽出手法の検討
3. 学会等名 日本音響学会2018年春季研究発表会講演論文集
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

神戸大学研究者紹介システム  
<http://kuid.ofc.kobe-u.ac.jp/InfoSearch/>  
研究室webページ  
<http://www.me.cs.scitec.kobe-u.ac.jp/~takigu/publications.html>

#### 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	高田 哲  (TAKADA Satoshi)  (10216658)	神戸大学・保健学研究科・名誉教授   (14501)	
研究分担者	陳 金輝  (CHEN Jinhui)  (50777810)	県立広島大学・地域創生学部・准教授   (25406)	
研究分担者	中井 靖  (NAKAI Yasushi)  (80462050)	宮崎大学・教育学部・准教授   (17601)	
研究分担者	山本 暁生  (YAMAMOTO Akio)  (30758842)	神戸大学・保健学研究科・特命助教   (14501)	

#### 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0件

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