

様式 C - 7 - 1

平成29年度科学研究費助成事業（科学研究費補助金）実績報告書（研究実績報告書）

所属研究機関名称		奈良先端科学技術大学院大学	機関番号	14603
研究代表者	部局	データ駆動型サイエンス創造センター		
	職	教授		
	氏名	中村 哲		

1. 研究種目名 基盤研究(S) 2. 課題番号 17H061013. 研究課題名 次世代音声翻訳の研究4. 研究期間 平成29年度～令和3年度 5. 領域番号・区分 -

6. 研究実績の概要

A)教師無し手法である独立低ランク行列分析に関して、そのスパース性を制御できるモデル(t分布・一般化ガウス分布)へ拡張し、スパース性付与が雑音抑圧能力向上に寄与することを実証。B)音響モデルと言語モデルをニューラルネットワークで統合的に内包した単語単位のEnd-to-End (Acoustic-to-Word) モデルを実現。MMMFを用いた雑音抑圧法の研究。C)統語的距離がある日英のEnd-to-Endアテンションベースの同時通訳システムを構築。半教師学習が可能なマシンスピーチチェーンを開発。D)多言語機械翻訳における原言語側のデータ欠落を特殊記号で補う手法を提案し、機械翻訳精度を改善。E)対話翻訳のデータセットを構築しIWSLT2017において公開。対話翻訳特有の現象について分析。

A)深層学習を用いて強調などの非言語情報に対応した音声翻訳手法を提案し、従来手法を大きく上回る翻訳結果を得た。B)音声波形モデリングに基づく統計的声質変換技術として、信号処理に基づく音声波形加工処理と深層学習に基づく音声波形生成処理に関する研究を進め、基盤技術を大いに改善。

口内部の不自然さを払拭すべく、歯形モデルの生成、発話時口内部のビデオシーケンスのデータベース化、リップシンクの口形状と同期した口内部映像の合成により、パッチベースのポアソンイメージエディティングの手法で実写クオリティの表情合成システムを開発。音声信号からリップシンクをリミテッドアニメ風を実現するVoiceAnimatorシステムを実現。

リアルタイムのユーザの状態推定を目指し脳波からの違和感を検出。意味違反および統語違反状態を意図的に作り出し、被験者に聞いてもらう実験で、有意な事象関連電位の出現および6割程度での機械学習モデルによる意味違反検出に成功。

学術シンポジウムの日英講演の同時通訳コーパス約5.5時間分を構築。

7. キーワード

音声翻訳

8. 現在までの進捗状況

区分 (2) おおむね順調に進展している。

理由
 常時音声認識において従来と全く異なる認識方式を考案したことなど、一部、当初の計画以上の進展があったが、それ以外についてはおおむね計画通りに研究が進んでいる。

3版

9. 今後の研究の推進方策

雑音抑制では、事前に用意された音源データを活用するため、独立低ランク行列分析における音源モデルパラメータ推定部をDNNによる推論に置き換える作業を進め、DNNとの融合を行う。スピーチチェーンの改良。英日間の機械翻訳における性能の検証。データ欠落を補う手法のさらなる性能向上。機械翻訳の基本方式の検討。対話翻訳特有の現象を考慮した翻訳手法を検討。翻訳の評価という観点で対話状態を用いるため対話状態付きの対訳データセットを構築。

感情翻訳など新たな特性をシステムに導入する。深層学習に基づく音声波形生成処理の改善に取り組むとともに、声質変換処理へと導入することで、高精度な声質変換システムの構築を目指す。Voice Conversion Challenge 2018に参加し、構築したシステムの性能を明らかにする。深層学習ベースの音声画像翻訳の手法に取り組む。違和感検出モデルのさらなる改良、および違和感検出精度の向上を目指す。また同時通訳作業中の聴取負荷の測定実験を始める。大規模な同時通訳コーパス構築のための体制を整備し100時間を超えるコーパス整備を行う。

10. 研究発表（平成29年度の研究成果）

〔雑誌論文〕 計4件（うち査読付論文 4件 / うち国際共著論文 2件 / うちオープンアクセス 4件）

1. 著者名 Quoc Truong Do, Tomoki Toda, Graham Neubig, Sakriani Sakti, and Satoshi Nakamura	4. 巻 25
2. 論文標題 Preserving Word-Level Emphasis in Speech-to-Speech Translation	5. 発行年 2017年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech and Language Processing	6. 最初と最後の頁 544-556
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TASLP.2016.2643280	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 Hiroki Tanaka, Hiroyoshi Adachi, Norimichi Ukita, Manabu Ikeda, Hiroaki Kazui, Takashi Kudo, Satoshi Nakamura	4. 巻 5
2. 論文標題 Detecting Dementia through Interactive Computer Avatars	5. 発行年 2017年
3. 雑誌名 IEEE Journal of Translational Engineering in Health and Medicine	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） 10.1109/JTEHM.2017.2752152	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 R.Duan, T.Kawahara, M.Dantsuji, and J.Zhang.	4. 巻 E100-D
2. 論文標題 Articulatory modeling for pronunciation error detection without non-native training data based on DNN transfer learning.	5. 発行年 2017年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 2174--2182
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transinf.2017EDP7019	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 Saito, Shinnosuke Takamichi, and Hiroshi Saruwatari	4. 巻 26
2. 論文標題 Statistical Parametric Speech Synthesis Incorporating Generative Adversarial Networks	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 84--96
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TASLP.2017.2761547	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47件(うち招待講演 3件/うち国際学会 21件)

1. 発表者名 角野隼斗, 北村大地, 高宗典玄, 高道慎之介, 猿渡洋, 小野順貴
2. 発表標題 独立深層学習行列分析に基づく多チャンネル音源分離
3. 学会等名 日本音響学会 2018年春季研究発表会
4. 発表年 2018年

1. 発表者名 Takatomo Kano, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Structured-Based Curriculum Learning for End-to-End English-Japanese Speech Translation
3. 学会等名 Interspeech 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 Quoc Truong Do, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Toward Expressive Speech Translation: A Unified Seq-to-Seq LSTMs Approach for Translating Words and Emphasis
3. 学会等名 Interspeech 2017 (国際学会)
4. 発表年 2017年

3版

1. 発表者名 Hiroki Watanabe, Hiroki Tanaka, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Subject-independent Classification of Japanese Spoken Sentences by Multiple Frequency Bands Phase Pattern of EEG Response during Speech Perception
3. 学会等名 Interspeech 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 Michael Heck, Masayuki Suzuki, Takashi Fukuda, Gakuto Kurata, Satoshi Nakamura
2. 発表標題 Ensembles of Multi-scale VGG Acoustic Models
3. 学会等名 Interspeech 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 K. Mukaihara, S. Sakti, S. Nakamura
2. 発表標題 Recognizing Emotionally Coloured Dialogue Speech using Speaker-Adapted DNN-CNN Bottleneck Features
3. 学会等名 SPECOM 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 M. Heck, S. Sakti, S. Nakamura
2. 発表標題 Feature Optimized DPGMM Clustering for Unsupervised Subword Modeling: A Contribution to ZEROSPEECH 2017
3. 学会等名 ASRU 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 A. Tjandra, S. Sakti, S. Nakamura
2. 発表標題 Attention-based Wav2Text with Feature Transfer
3. 学会等名 ASRU 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 A. Tjandra, S. Sakti, S. Nakamura
2. 発表標題 Listening while Speaking: Speech Chain by Deep Learning
3. 学会等名 ASRU 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 A. Tjandra, S. Sakti, S. Nakamura
2. 発表標題 Local Monotonic Attention Mechanism for End-to-end Speech and Language Processing
3. 学会等名 IJCNLP 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 Yusuke Oda, Philip Arthur, Graham Neubig, Koichiro Yoshino, Satoshi Nakamura
2. 発表標題 Neural Machine Translation via Binary Code Prediction
3. 学会等名 55th Annual Meeting of the Association for Computational Linguistics (ACL) (Long Papers) (国際学会)
4. 発表年 2017年

3版

1. 発表者名 A. Tjandra, S. Sakti, S. Nakamura
2. 発表標題 End-to-end Speech Recognition with Local Monotonic Attention
3. 学会等名 NIPS Workshop (国際学会)
4. 発表年 2017年

1. 発表者名 柳田 智也, S. Sakti, 中村 哲
2. 発表標題 日本語インクリメンタル音声合成システム実装のための言語特徴の検討
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 2017年

1. 発表者名 森 巧磨, Andros Tjandra, Sakriani Sakti, 中村 哲
2. 発表標題 テンソルトレイン分解によるEnd-to-End自動音声認識モデルの圧縮
3. 学会等名 情報処理学会 自然言語処理研究会
4. 発表年 2017年

1. 発表者名 Masahiro Honda, Hiroki Tanaka, Sakti Sakriani, Satoshi Nakamura
2. 発表標題 Detecting Suppression of Negative Emotion by Time Series Change of Cerebral Blood Flow using fNIRS
3. 学会等名 IEEE International Conference on Biomedical and Health Informatics (BHI) (国際学会)
4. 発表年 2018年

1. 発表者名 Naoto Terasawa, Hiroki Tanaka, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Tracking Liking State in Brain Activity while Watching Multiple Movies
3. 学会等名 19th ACM International Conference on Multimodal Interaction (ICMI'17) (国際学会)
4. 発表年 2017年

1. 発表者名 Naoto Terasawa, Hiroki Tanaka, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 EEG-based Emotional State Tracking during Watching Movie considering Self-Assessment Manikin
3. 学会等名 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC2017) (国際学会)
4. 発表年 2017年

1. 発表者名 Johanes Effendi, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Creation of a Multi-paraphrase Corpus based on Various Elementary Operations
3. 学会等名 Oriental COCOSDA 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 Michael Heck, Masayuki Suzuki, Takashi Fukuda, Gakuto Kurata, Satoshi Nakamura
2. 発表標題 Distilling Knowledge from a Multi-scale deep CNN Ensemble for Robust and Light-weight Acoustic Modeling
3. 学会等名 第120回音声言語情報処理研究会 (IPSJ SIG-SLP)
4. 発表年 2018年

3版

1. 発表者名 長村佳歩, 叶高朋, SakrianiSakti, 須藤克仁, 中村哲
2. 発表標題 音声認識単語仮説の曖昧性を考慮する ニューラル機械翻訳
3. 学会等名 言語処理学会第24回年次大会(NLP2018)
4. 発表年 2018年

1. 発表者名 西村優汰, 須藤克仁, 中村哲
2. 発表標題 原言語側の欠落を考慮したMulti-Source NMT
3. 学会等名 言語処理学会第24回年次大会(NLP2018)
4. 発表年 2018年

1. 発表者名 田中 宏季, 渡部 宏樹, 真木 勇人, Sakriani Sakti, 中村 哲
2. 発表標題 音声文聴取時における意味違反が生じた際の脳波自動判別
3. 学会等名 電子情報通信学会技術研究報告ヒューマン情報処理研究会 (HIP)
4. 発表年 2017年

1. 発表者名 宇城 毅穣, 田中 宏季, 足立 浩祥, 數井 裕光, 池田 学, 工藤 喬, 中村 哲
2. 発表標題 エージェントによる非定型質問への応答からの認知症検出
3. 学会等名 IPSJ SIG
4. 発表年 2018年

1. 発表者名 真木 勇人, Sakriani Sakti, 田中 宏季, 中村 哲
2. 発表標題 EEGを用いた合成音声に対する体感品質予想
3. 学会等名 電子情報通信学会MEとバイオサイバネティックス研究会 (MBE)
4. 発表年 2018年

1. 発表者名 真木 勇人, 田中 宏季, Sakriani Sakti, 中村 哲
2. 発表標題 電極配置のグラフ構造を利用したテンソル分解による単一試行EEG解析
3. 学会等名 電子情報通信学会MEとバイオサイバネティックス研究会 (MBE)
4. 発表年 2018年

1. 発表者名 田中 宏季, 寺澤 直人, 本田 将大, 真木 勇人, サクリアニ サクティ
2. 発表標題 生体信号からの感情コンピューティングと自閉症支援
3. 学会等名 第13回日本感性工学会春季大会
4. 発表年 2018年

1. 発表者名 Michael Heck, Masayuki Suzuki, Takashi Fukuda, Gakuto Kurata and Satoshi Nakamura
2. 発表標題 Knowledge Distillation from a Multi-scale VGG Ensemble for Acoustic Modeling
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

3版

1. 発表者名 Nurul Lubis, Sakriani Sakti, Koichiro Yoshino and Satoshi Nakamura
2. 発表標題 Dialogue Modeling for Eliciting Positive Emotion
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 Quoc Truong Do, Sakriani Sakti and Satoshi Nakamura
2. 発表標題 Joint Translation of Words and Emphasis in Speech-to-Speech Translation using Sequence-to-Sequence Models
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 叶 高朋, Sakriani Sakti, 中村 哲
2. 発表標題 カリキュラムラーニングを用いた日英直接翻訳システムの提案
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 Andros Tjandra, Sakriani Sakti and Satoshi Nakamura
2. 発表標題 Tensor Train based RNN Compression for Polyphonic Music Modelling
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 中村 哲
2. 発表標題 音声翻訳研究のこれから
3. 学会等名 日本音響学会2017年秋季研究発表会（招待講演）
4. 発表年 2017年

1. 発表者名 M.Mimura, S.Sakai, and T.Kawahara
2. 発表標題 Cross-domain speech recognition using nonparallel corpora with cycle-consistent adversarial networks.
3. 学会等名 IEEE Workshop Automatic Speech Recognition & Understanding (ASRU) (国際学会)
4. 発表年 2017年

1. 発表者名 T.Kawahara
2. 発表標題 Automatic meeting transcription system for the Japanese Parliament (Diet).
3. 学会等名 APSIPA ASC (招待講演) (国際学会)
4. 発表年 2017年

1. 発表者名 T.Kawahara
2. 発表標題 Modeling difficulties of second language learners using speech technology.
3. 学会等名 Seoul International Conference on Speech Sciences (SICSS) (招待講演) (国際学会)
4. 発表年 2017年

3版

1. 発表者名 M.Wake, Y.Bando, M.Mimura, K.Itoyama, K.Yoshii, and T.Kawahara.
2. 発表標題 Semi-blind speech enhancement based on recurrent neural network for source separation and dereverberation.
3. 学会等名 IEEE Machine Learning for Signal Processing Workshop (MLSP) (国際学会)
4. 発表年 2017年

1. 発表者名 M.Mimura, Y.Bando, K.Shimada, S.Sakai, K.Yoshii, and T.Kawahara.
2. 発表標題 Combined multi-channel NMF-based robust beamforming for noisy speech recognition.
3. 学会等名 INTERSPEECH (国際学会)
4. 発表年 2017年

1. 発表者名 H.Inaguma, K.Inoue, M.Mimura, and T.Kawahara.
2. 発表標題 Social signal detection in spontaneous dialogue using bidirectional LSTM-CTC.
3. 学会等名 INTERSPEECH (国際学会)
4. 発表年 2017年

1. 発表者名 島田一希, 坂東宜昭, 三村正人, 糸山克寿, 吉井和佳, 河原達也.
2. 発表標題 マルチチャンネル非負値行列因子分解に基づく ビームフォーミングを用いた雑音環境下音声認識.
3. 学会等名 電子情報通信学会SP
4. 発表年 2018年

1. 発表者名 上乃聖, 稲熊寛文, 三村正人, 河原達也.
2. 発表標題 CTCによる文字単位のモデルを併用したattentionによる単語単位の end-to-end音声認識.
3. 学会等名 情報処理学会SIG-SLP
4. 発表年 2018年

1. 発表者名 島田一希, 坂東宜昭, 三村正人, 糸山克寿, 吉井和佳, 河原達也.
2. 発表標題 雑音環境下音声認識のための多チャネル非負値行列因子分解に基づく教師なしピームフォーマ.
3. 学会等名 電子情報通信学会SP
4. 発表年 2017年

1. 発表者名 和気雅弥, 坂東宜昭, 三村正人, 糸山克寿, 吉井和佳, 河原達也.
2. 発表標題 再帰型ニューラルネットワークを用いたセミブラインド音声分離・強調.
3. 学会等名 電子情報通信学会SP
4. 発表年 2017年

1. 発表者名 坂東宜昭, 三村正人, 糸山克寿, 吉井和佳, 河原達也.
2. 発表標題 深層生成モデルを事前分布に用いた教師なし音声強調.
3. 学会等名 電子情報通信学会SP
4. 発表年 2017年

3 版

1. 発表者名 稲熊寛文, 井上昂治, 三村正人, 河原達也.
2. 発表標題 End-to-endモデルによるsocial signals検出および音声認識との統合.
3. 学会等名 情報処理学会SIG-SLP
4. 発表年 2017年

1. 発表者名 北村大地, 角野隼斗, 高宗典玄, 高道慎之介, 猿渡洋, 小野順貴
2. 発表標題 独立深層学習行列分析に基づく多チャンネル音源分離の実験的評価
3. 学会等名 電子情報通信学会技術研究報告音声研究会 (SP)
4. 発表年 2018年

1. 発表者名 P.L. Tobing, Y.-C. Wu, T. Hayashi, K. Kobayashi, T. Toda
2. 発表標題 Development of NU voice conversion system 2018
3. 学会等名 電子情報通信学会技術研究報告音声研究会 (SP)
4. 発表年 2018年

1. 発表者名 Y.-C. Wu, P.L. Tobing, T. Hayashi, K. Kobayashi, T. Toda
2. 発表標題 Development of NU non-parallel voice conversion system 2018
3. 学会等名 電子情報通信学会技術研究報告音声研究会 (SP)
4. 発表年 2018年

〔図書〕 計0件

1 1 . 研究成果による産業財産権の出願・取得状況

計0件（うち出願0件 / うち取得0件）

1 2 . 科研費を使用して開催した国際研究集会

計0件

1 3 . 本研究に関連して実施した国際共同研究の実施状況

-

1 4 . 備考

-