

人間と情報環境の共生インタラクション基盤技術の創出と展開  
2018年度採択研究代表者

2020年度 年次報告書
-----------------

山岸 順一

情報・システム研究機構国立情報学研究所コンテンツ科学研究系  
教授

VoicePersonae: 声のアイデンティティクローニングと保護

## § 1. 研究成果の概要

VoicePersonae プロジェクトでは、声のアイデンティティに関する分野の壁を取り除き、①アイデンティティに関する機械学習技術を高精度化すると同時に、②生体認証の安全性と頑健性を高め、③プライバシー保護に関する新しい技術を実現する。また音声変換と生体検知、匿名化と再識別化と言った目的が相反する技術をどちらも加速させる敵対的競争型研究を実施する。更に他のモダリティ情報へ研究成果を適用し、アイデンティティの活用と保護を両立させる技術を確立させることを目指している。

本年度は、これまで取り組んできた、音声合成、声質変換、話者認識、ライブネス検知の融合および発展に関する多くの学術的成果をジャーナル論文としてまとめ上げると同時に、研究対象を広げることも行なった。その1つが、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話者の属性情報を保ったまま、音声の個人性を変えることを目的とする「話者匿名化」の研究である。この新しい基礎研究課題の発展のため、仏 Avignon 大、Eurecom 研究所、Inria 研究所、および NII が協力し、話者匿名化法の相互比較のために利用すべき音声データベース、評価セット、ベースライン、そして、評価方法を規定し、無償で一般公開した。

これに加え、聞き手側の環境に存在する雑音に負けないよう音声を予め変換する「音声明瞭性強調」の研究も新たに開始した。単純な教師あり学習では解決できないチャレンジングな課題であるが、複雑で微分不可能な音声明瞭性の指標をニューラルネットワークの識別器に近似させ、そして、その近似指標に基づき、不明瞭な音声を聴きやすい音声に自動変換する枠組を提案し、音声の明瞭性が劇的に向上することを実験から示した。

このほか、ディープフェイク動画検出技術の社会還元に向けた取り組みや、顔認証システムの脆弱性の調査も行い、其々成果を挙げた。

## § 2. 研究実施体制

### (1) NII グループ

- ① 研究代表者: 国立情報学研究所コンテンツ科学研究系 教授 山岸順一
- ② 研究項目
  - ・声のアイデンティティのモデル化に関する理論的統合
  - ・音声合成および声質変換に関する研究
  - ・話者認識の安全性と頑健性の向上に関する研究
  - ・音声のプライバシー保護に関する研究
  - ・他の生体情報におけるライブネス検出の研究

### 【代表的な原著論文情報】

1. Li Haoyu, Szu-Wei Fu, Yu Tsao, Junichi Yamagishi, "iMetricGAN: Intelligibility Enhancement for Speech-in-Noise using Generative Adversarial Network-based Metric Learning", Interspeech 2020, 2020 年 10 月
2. Xin Wang, Junichi Yamagishi, Massimiliano Todisco, Héctor Delgado, Andreas Nautsch, Nicholas Evans, Md Sahidullah, Ville Vestman, Tomi Kinnunen, Kong Aik Lee, Lauri Juvela, Paavo Alku, Yu-Huai Peng, Hsin-Te Hwang, Yu Tsao, Hsin-Min Wang, Sébastien Le Maguer, Markus Becker, Fergus Henderson, Rob Clark, Yu Zhang, Quan Wang, Ye Jia, Kai Onuma, Koji Mushika, Takashi Kaneda, Yuan Jiang, Li-Juan Liu, Yi-Chiao Wu, Wen-Chin Huang, Tomoki Toda, Kou Tanaka, Hirokazu Kameoka, Ingmar Steiner, Driss Matrouf, Jean-François Bonastre, Avashna Govender, Srikanth Ronanki, Jing-Xuan Zhang, Zhen-Hua Ling, "ASVspoof 2019: a large-scale public database of synthesized, converted and replayed speech", Computer Speech & Language, 64, 2020 年 11 月
3. Andreas Nautsch, Xin Wang, Nicholas Evans, Tomi Kinnunen, Ville Vestman, Massimiliano Todisco, Hector Delgado, Md Sahidullah, Junichi Yamagishi, Kong Aik Lee, "ASVspoof 2019: spoofing countermeasures for the detection of synthesized, converted and replayed speech", IEEE Transactions on Biometrics, Behavior, and Identity Science, 3,(2), 252-265, 2021 年 4 月
4. Natalia Tomashenko and Brij Mohan Lal Srivastava and Xin Wang and Emmanuel Vincent and Andreas Nautsch and Junichi Yamagishi and Nicholas Evans and Jose Patino and Jean-François Bonastre and Paul-Gauthier Noé and Massimiliano Todisco, "Introducing the VoicePrivacy Initiative", Interspeech 2020, 2020 年 10 月
5. Yi Zhao, Wen-Chin Huang, Xiaohai Tian, Junichi Yamagishi, Rohan Kumar Das, Tomi Kinnunen, Zhenhua Ling, Tomoki Toda, "Voice Conversion Challenge 2020: Intra-lingual semi-parallel and cross-lingual voice conversion", ISCA Joint Workshop for the Blizzard Challenge and Voice Conversion Challenge 2020, 2020 年 10 月
6. Huy H. Nguyen, Junichi Yamagishi, Isao Echizen, Sebastien Marcel, "Generating Master Faces for Use in Performing Wolf Attacks on Face Recognition Systems", INTERNATIONAL JOINT CONFERENCE ON BIOMETRICS (IJCB 2020), 2020 年 9 月