G-1

骨導音声と気導音声のスタイル変換に関する研究

A Study on Style Conversion between Bone-conducted and Air-conducted Speech

○村田京平[1]，関弘翔[2]，細野裕行[2]
*Kyohei Murata[1], Hiroto Seki[2], Hiroyuki Hosono[2]

Abstract: In this study, we proposed CycleGAN, which is a style conversion of bone-conducted speech into air-conduducted speech.

１．まえがき

近年，工場や建設現場などの高騒音環境下かつ防護具使用時における音声伝送の手段として，背景雑音の低減性に優れた骨導マイクロホンの利用が提案されている．しかし，骨導マイクロホンは気導マイクロホンに比べ収音できる帯域が狭く高周波成分が減衰し特性も異なることから音響ミスマッチが生じ，気導マイクロホンを想定した通常の音響モデルでは音声認識性能が低下する[1].

本研究では，音響ミスマッチをスタイルの違いと捉え，骨導マイクロホンで録音した音声(骨導音声)を気導マイクロホンで録音した音声(気導音声)にスタイル変換することで，骨導音声の認識率の向上が可能になると考え手法を提案する．

２．先行研究とその課題

骨導音声の認識率の向上を行うことを目的とした先行研究に藤岡らの研究がある[2]. 藤岡らの研究では，回帰問題と捉え Convolutional Neural Network(CNN)を用いて骨導音声を気導音声の様に補正する手法が提案された．この手法では学習に使用した話者に対する骨導音声の認識率は，36.3%から 76.3%に向上した．しかし，未知の話者に対する認識率は学習に使用した話者のように大きく向上しないこと，また，CNN の回帰学習による手法では同一話者により同時に録音した各音声のペアデータが必要になるという課題が残された．よって本研究では，2 つの対になっていない(アンペア)のデータを基に，Cycle-ConsistentAdversarial Networks (CycleGAN)[3]を用いて骨導音声を気導音声の様に変換する手法を提案する．

３．提案手法

まず，CycleGAN を用いてアンペアのデータから骨導音声と気導音声相互のスタイル変換の学習を行う．CycleGAN は Generative Adversarial Networks(GAN)の一種で，偽物のデータを生成する Generator と，偽物か本物かを識別する Discriminator を 2 つずつ用いて構成される．それらを敵対的に学習させ，2 つの画像データセット同士の domain(分野，領域)の関係を学習してスタイル変換を行うことで，アンペアの画像データセットから自然な変換を行うことが可能となる．

本研究では，気導音声は音声コーパスである JVS corpus[4]を利用し，骨導音声は気導音声と同様の発話内容を，骨導マイクロホンで録音し利用する．各音声をメルスペクトログラム化し，それぞれ Air domain, Bone domain とする．CycleGAN を用いてこの domain 間でスタイル変換の学習を行う．骨導音声を気導音声に変換する Generator を用いて，変換後のメルスペクトログラムから音声を再構成し，Google 社が提供する音声認識エンジン「Cloud Speech to Text」で認識させ音声認識の精度を比較する．提案手法の模式図を Figure 1 に示す．

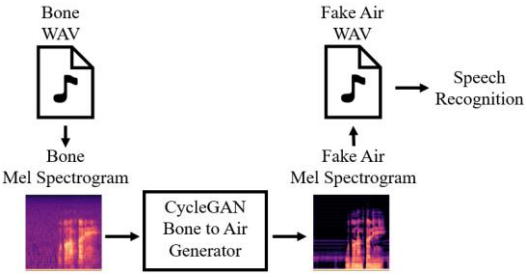

Figure 1. Proposed system schematic diagram

４．まとめ

本研究では骨導音声を気導音声にスタイル変換する CycleGAN を提案した．

参考文献

[1] 林升柯，他：「複数の装着型マイクを用いた多人数会話音声認識に関する検討」，情報処理学会，2016
[2] 藤岡・関・細野：「骨導音声の認識率向上の検討」，2020 年電子情報通信学会総合大会，D-14-5，pp.112
[3] Jun-Yan Zhu, et al. : “Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks,” in Proc.ICCV 2017，pp.2223-2232 (2017)
[4] S. Takamichi, et al. : “JVS corpus: free Japanese multi-speaker voice corpus,” arXiv preprint, 1908.06248 (2019)

１：日大理工・学部・情報　２：日大理工・教員・情報