

AIするディープラーニング AI Love Deep Learning

メンバー：濱口和希* 白鳥孝幸 山田大貴 齋藤匠 三浦幸泰 坂下夏槻 和田孝喜 川村一世 松元健 *リーダー

担当教員：竹之内高志 香取勇一 寺沢憲吾 片桐恭弘 富永敦子

Member : Kazuki Hamaguchi* / Takayuki Shiratori / Daiki Yamada / Takumi Saitou / Kouta Miura / Katsuki Sakashita / Kouki Wada / Issei Kawamura / Takeru Matsumoto * : Leader

Supervisor : Takashi Takenouchi / Yuichi Katori / Kengo Terasawa / Yasuhiro Katagiri / Atsuko Tominaga

概要 Overview

ディープラーニングとは What's Deep Learning

- ・人口の脳神経回路網(Artificial Neural Network : ANN)を元に作られた機械学習手法

Machine learning method based on Artificial Neural Network

- より層を深く(ディープに)したANNを用いる

Deep Learning uses ANN with deeper layer

- より高精度な学習が可能

Deep Learning can learn more accurately

- ・従来の手法では不可能だった課題を解決できる

Deep Learning can solve problems could not be done with previous technology

- 強化学習 Reinforcement Learning
- 自然言語処理 Natural Language Processing
- 画像認識 Image Recognition
- 生成モデル Generative Model

目標 Purpose

- ・ディープラーニングを使用し応用技術の開拓に挑戦する

We challenge to the development of applied technology by deep learning



Group A : 声質変換

Voice Conversion

Group B : 姿勢推定

Pose Estimation

声質変換 Voice Conversion

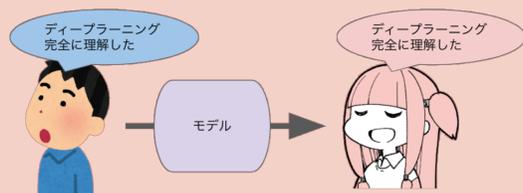
背景と目的 Background and Purpose

- ・個人の動画配信、バーチャルYouTuberの活動において、声質変換の需要が高まっている

In personal video streaming and virtual YouTuber activities, the demand of voice conversion is increasing

- ・ディープラーニングによって、自身の音声を、VOICEROIDである琴葉茜[1]の音声に変換する

We'll try to convert particular voice to voice of Kotonoha Akane by using DeepLearning



[1] 「VOICEROID2 琴葉茜・葵」 <https://www.ah-soft.com/voiceroid/kotonoha/> (参照2018-11-30).

各モデル Each models

- ・1段階目のモデル First model
- ・2段階目のモデル Second model

- 自身の音声を、琴葉茜の音声(低音質)に変換

It converts particular voice to Akane

- 低音質な琴葉茜の音声を、高音質な琴葉茜の音声に変換

It converts low-quality Akane voice to high-quality Akane Voice

- 学習にはパラレルデータ*が必要

It is needed parallel datas to learn

- 学習には大量の琴葉茜の音声が必要

It is needed a lot of voice datas of Akane to learn



*パラレルデータ：入力話者と出力話者について同時に同じ内容を発話した音声データ

姿勢推定 Pose Estimation

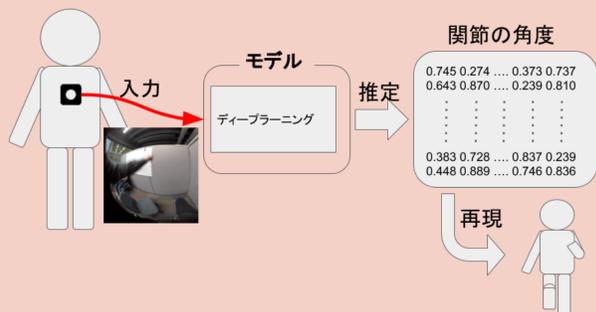
背景と目的 Background and Purpose

- ・現状、モーションキャプチャーには様々な機材や広いスペースを必要とする

Now, motion capture needs various equipments and large space

- ・第三者視点のカメラも大掛かりな機材も必要ないモーションキャプチャの手法を提案

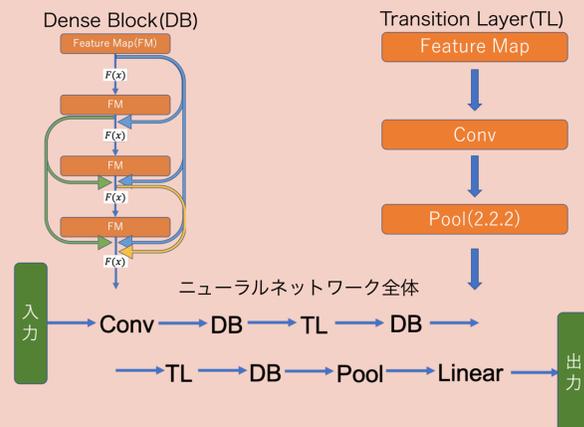
We propose a motion capture technique without neither large equipments nor third person camera.



実験と結果 Experiments and Results

- ・学習には72層のDenseNetを使用した

We used DenseNet-72



- ・十分な精度を持つモデルを作ることはできなかった

We could not construct good estimator