Crossmodal action unit sequence estimation and image-to-image mapping

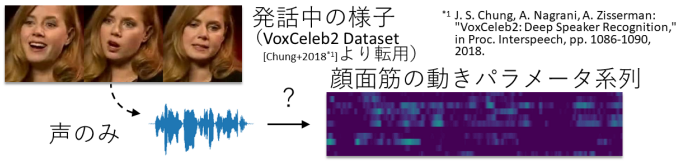
17 クロスモーダルアクションユニット系列推定と顔画像変換

声で顔画像の表情を動かす

どんな研究	音声には発話内容に相当する言語情報だけでなく、感情表現やムードに相当する非言語情報が含まれ、音声対話において重要な役割を担っています。本研究では、音声の非言語情報は話者の顔表情に表出されていると仮定し、**音声のみから話者のアクションユニット（顔面筋パラメータ）を推定**することを初めて試みた研究です。
どこが凄い	これまで音声のみからアクションユニットを推定する試みはなく、**どの程度の精度を達成できるかは未知数でしたが、本研究ではこれを初めて明らかに**しました。また、音声から推定したアクションユニットと画像変換器を用いることで、声に合わせて静止顔画像の表情を動かすシステムを実装し、**声の表情や雰囲気を可視化**することを可能にしました。
めざす未来	感情表現やムードは、従来、主観に基づく大まかなラベルにより記号的に扱われることが主流でした。これに対し、アクションユニットは**感情表現やムードを細やかに表現する連続量として好適**であり、本研究で音声からアクションユニットを推定できることを示しました。今後、顔表情に合った音声合成、音声に合った顔画像生成など、**音声と顔画像を同時利用した様々な応用技術が拓ける**と期待しています。

音声から顔の動きを推定できるか？

✓ 声から顔面筋の動きを予測できれば・・・



*1 J. S. Chung, A. Nagrani, A. Zisserman: "VoxCeleb2: Deep Speaker Recognition," in Proc. Interspeech, pp. 1086-1090, 2018.

・予測結果を顔画像に転写することで音声の非言語情報を可視化可能に
・音声の非言語情報を表現する有用な特徴量として利用可能に
（音声合成や音声変換に応用可能）

✓ どれほど難しいのか？そのようなことが可能なのか？
➡ 本研究はこの問いに答えるもの

深層学習と音声付き動画像を用いた解法

✓ 顔面筋の動きを表したパラメータとして
アクションユニット(Action Unit; AU)[*2]に着目

*2 **顔の表情動作の最小単位で、眉毛・瞼・口・唇の動きを数値化したもの**

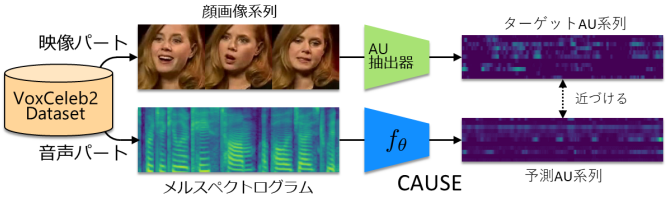
✓ 声からAUを予測するニューラルネットワークを学習
「Crossmodal Action Unit Sequence Estimator (CAUSE)」

アプローチ

大量の音声付き動画を用いて
- 映像パートからOpenFace[*3]により抽出されるAU系列
- 音声パートからCAUSEにより予測されるAU系列

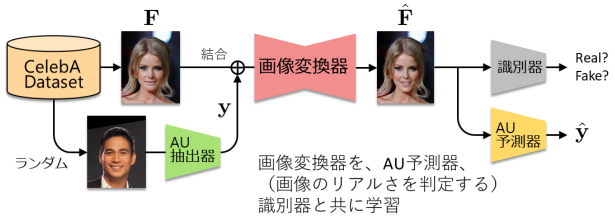
が近くなるようにCAUSEを学習

*3オープンソースの顔動作解析ツールキット



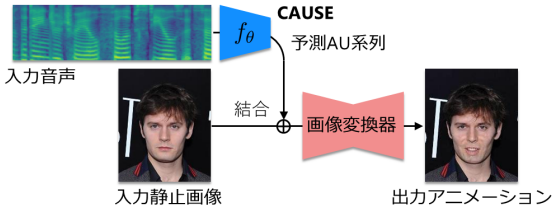
クロスモーダル顔画像制御

✓ 顔画像変換器の学習：GANimation [Pumarola+2018]



画像変換器を、AU予測器、（画像のリアルさを判定する）識別器と共に学習

✓ 声から予測したAU系列を用いて静止顔画像を変換



➡ 声を用いて顔画像の表情を制御することが可能

※ 顔画像はいずれもCelebA Dataset [Liu+2015*4]のものを使用

*4 Z. Liu, P. Luo, X. Wang, X. Tang: "Deep Learning Face Attributes in the Wild," in Proc. ICCV, pp. 3730-3738, 2015.

顔画像制御実験

その他の例はこちら

✓ 同一音声で異なる顔画像を制御した例

✓ AUの代わりに感情状態（平常、喜び、驚き、悲しみ、怒り、嫌悪、恐怖、軽蔑）の確率ベクトルを介して制御を行った場合に比べ、自然な顔アニメーションを生成できることを確認

関連文献

[1] H. Kameoka, T. Kaneko, S. Seki, K. Tanaka, “CAUSE: Crossmodal action unit sequence estimation from speech,” submitted to The 23rd Annual Conference of the International Speech Communication Association (Interspeech 2022).

連絡先

亀岡 弘和（Hirokazu Kameoka） メディア情報研究部 メディア認識研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp