

16

あなた専用のお手本映像で上達支援

～深層学習によるメディア生成の可能性～



どんな研究

「したいことがあるけど、やり方が分からない…」そのような時に、ユーザが与えた音声や画像などのメディア情報を元に**具体的な指示や理想像を提示**するシステムの実現を目指しています。本展示では、このシステム実現のために行った二つの取組について紹介します。

どこが凄い

従来のフィードバックシステムは、ルールベースのものが主で、ユーザの多様な入力状態に対して対応することは困難でした。これに対し、本研究では、**深層学習をベースに新たな情報伝播方法・ネットワーク構造**を考案し、「あなた専用のお手本」を提示することを可能にしました。

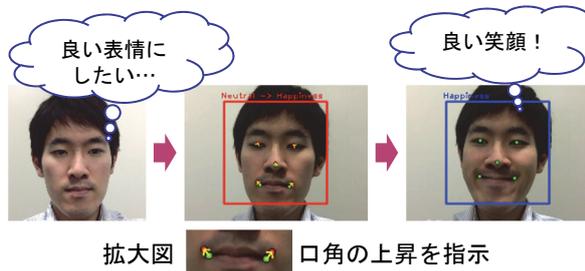
めざす未来

本研究の鍵は、システムが**メディア生成を通していかにユーザに寄り添えるか**ということです。これは、上達支援だけにとどまらず、ユーザが思い描くメディアを生成するために必要な技術です。将来的にはユーザのあらゆる要望に応えられる**メディア生成技術の確立**を目指しています。

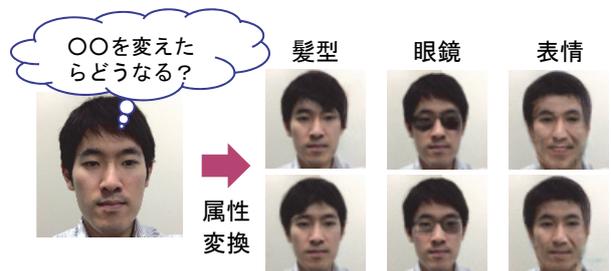
「あなた専用のお手本映像」の生成（例：顔の表情改善・顔の属性変換の場合）



具体的な指示の提示（例：顔の表情改善の場合）

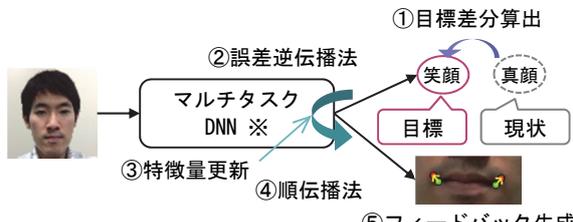


理想像の提示（例：顔の属性変換の場合）



Feedback Propagation [1]

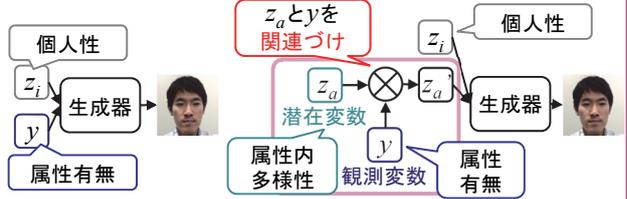
フィードバック生成のための**情報伝播方法**を提案



※表情認識とキーポイント検出について学習済みモデルを利用

Conditional Filtered GAN [2]

属性内多様性の学習のための**ネットワーク構造**を提案



属性の有無変換のみ
従来法：CGAN

属性の**多様な変換**が可能
提案法：CFGAN

関連文献

- [1] T. Kaneko, K. Hiramatsu, K. Kashino, "Adaptive visual feedback generation for facial expression improvement with multi-task deep neural networks," in *Proc. The 24th ACM International Conference on Multimedia (ACMMM)*, pp. 327-331, 2016.
- [2] T. Kaneko, K. Hiramatsu, K. Kashino, "Generative attribute controller with conditional filtered generative adversarial networks," in *Proc. The 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017 (to appear).
- [3] T. Kaneko, H. Kameoka, N. Hojo, Y. Ijima, K. Hiramatsu, K. Kashino, "Generative adversarial network-based postfilter for statistical parametric speech synthesis," in *Proc. The 42nd IEEE International Conference on Acoustic, Speech and Signal Processing*, 2017.

連絡先

金子 卓弘 (Takuhiko Kaneko) メディア情報研究部 メディア認識研究グループ
E-mail: kaneko.takuhiro(at)lab.ntt.co.jp