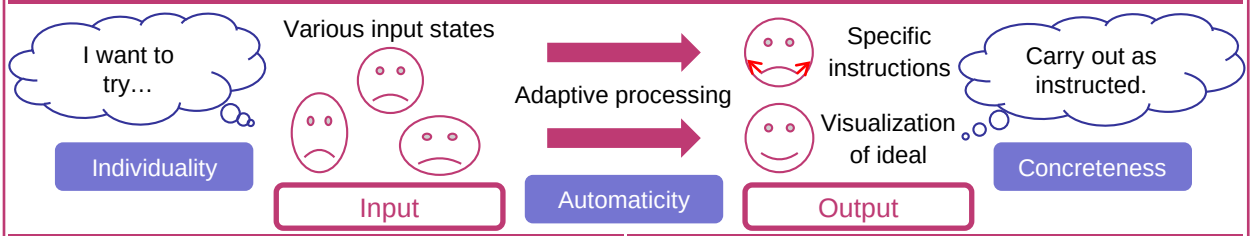
16 Generative personal assistance

- Deep learning opens the way to innovative media generation -

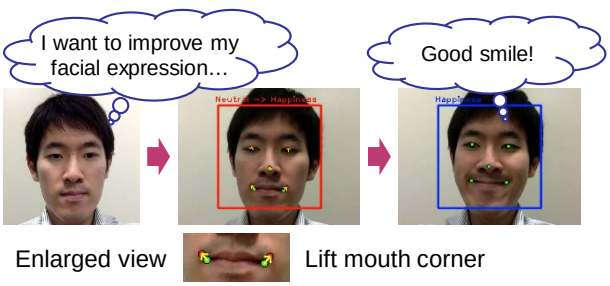
Abstract

We aim to develop a system that can give feedback or instructions to a person who wishes to better do something or do new things. Unlike the existing personal assistance methods based on manually defined rules, our goal is to develop a system with the following advantages: (1) individuality, meaning that the system output must be suitable for individual persons; (2) concreteness, so that the instructions are concrete enough for users to easily understand; and (3) automaticity, so that the system performs the above process automatically. To this end, we propose two kinds of learning-based (specifically, deep learning-based) approaches: One is a novel information propagation method to generate appropriate feedback according to inputs, and the other is a novel network architecture to represent attribute variations. We believe that these approaches will also lead to a generic media generation technique that will meet a variety of demands in the near future.

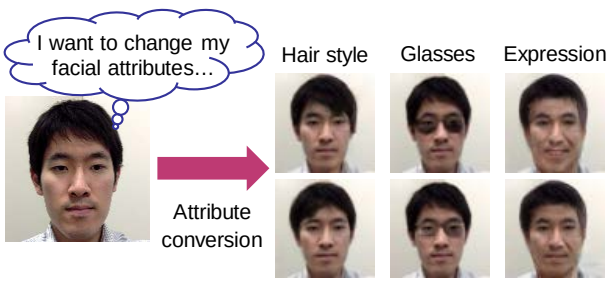
Generative personal assistance (e.g., facial expression improvement, face attribute conversion)



Specific instructions

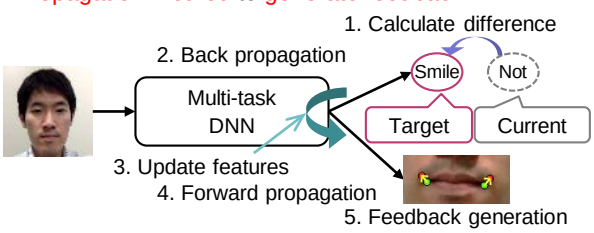


Visualization of ideal



Feedback Propagation [1]

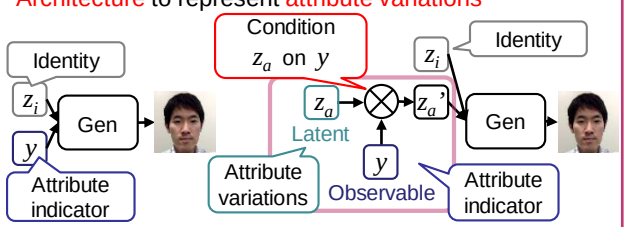
Propagation method to generate feedback



Use models learned for facial expression classification and face keypoint detection

Conditional Filtered GAN [2]

Architecture to represent attribute variations



ON or OFF control — Variation control

Previous: CGAN — Proposed: CFGAN

Reference

[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.

Contact

Takuhiro Kaneko Recognition Research Group, Media Information Laboratory
Email : kaneko.takuhiro(at)lab.ntt.co.jp