

AI could make dodgy lip sync dubbing a thing of the past

August 17 2018, by Vicky Just



The system enables film editors to change the expression and head position of an actor after filming has finished. Credit: University of Bath

Researchers have developed a system using artificial intelligence that can edit the facial expressions of actors to accurately match dubbed voices, saving time and reducing costs for the film industry. It can also be used to correct gaze and head pose in video conferencing, and enables new possibilities for video postproduction and visual effects.

The technique was developed by an international team led by a group from the Max Planck Institute for Informatics and including researchers from the University of Bath, Technicolor, TU Munich and Stanford University. The work, called [Deep Video Portraits](#), was presented for the first time at the SIGGRAPH 2018 conference in Vancouver on 16th August.

Unlike previous methods that are focused on movements of the face interior only, Deep Video Portraits can also animate the whole face including eyes, eyebrows, and head position in videos, using controls known from computer graphics face animation. It can even synthesise a plausible static [video](#) background if the head is moved around.

Hyeongwoo Kim from the Max Planck Institute for Informatics explains: "It works by using model-based 3-D face performance capture to record the detailed movements of the eyebrows, mouth, nose, and head position of the dubbing actor in a video. It then transposes these movements onto the 'target' actor in the film to accurately sync the lips and facial movements with the new audio."

The research is currently at the proof-of-concept stage and is yet to work at real time, however the researchers anticipate the approach could make a real difference to the visual entertainment industry.

Professor Christian Theobalt, from the Max Planck Institute for Informatics, said: "Despite extensive post-production manipulation, dubbing films into foreign languages always presents a mismatch between the actor on screen and the dubbed voice.

"Our new Deep Video Portrait approach enables us to modify the appearance of a target actor by transferring head pose, [facial expressions](#), and eye motion with a high level of realism."

Co-author of the paper, Dr. Christian Richardt, from the University of Bath's motion capture research centre CAMERA, adds: "This technique could also be used for post-production in the [film industry](#) where computer graphics editing of faces is already widely used in today's feature films."

A great example is 'The Curious Case of Benjamin Button' where the face of Brad Pitt was replaced with a modified computer graphics version in nearly every frame of the movie. This work remains a very time-consuming process, often requiring many weeks of work by trained artists.

"Deep Video Portraits shows how such a visual effect could be created with less effort in the future. With our approach even the positioning of an actor's head and their facial expression could be easily edited to change camera angles or subtly change the framing of a scene to tell the story better."

In addition, this new approach can also be used in other applications, which the authors show on their project website, for instance in video and VR teleconferencing, where it can be used to correct gaze and head pose such that a more natural conversation setting is achieved. The software enables many new creative applications in visual media production, but the authors are also aware of the potential of misuse of modern video editing technology.

Dr. Michael Zollhöfer, from Stanford University, explains: "The media industry has been touching up photos with photo-editing software for many years, meaning most of us have learned to take what we see in photos with a pinch of salt. With ever improving video editing technology, we must also start being more critical about the video content we consume every day, especially if there is no proof of origin. We believe that the field of digital forensics should and will receive a lot

more attention in the future to develop approaches that can automatically prove the authenticity of a video clip. This will lead to ever better approaches that can spot such modifications even if we humans might not be able to spot them with our own eyes."

To address this, the research team is using the same technology to develop in tandem neural networks trained to detect synthetically generated or edited video at high precision to make it easier to spot forgeries. The authors have no plans to make the software publicly available but state that any software implementing the many creative use cases should include watermarking schemes to clearly mark modifications.

More information: richardt.name/publications/deep-video-portraits/

Provided by University of Bath

Citation: AI could make dodgy lip sync dubbing a thing of the past (2018, August 17) retrieved 19 April 2024 from <https://techxplore.com/news/2018-08-ai-dodgy-lip-sync-dubbing.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.