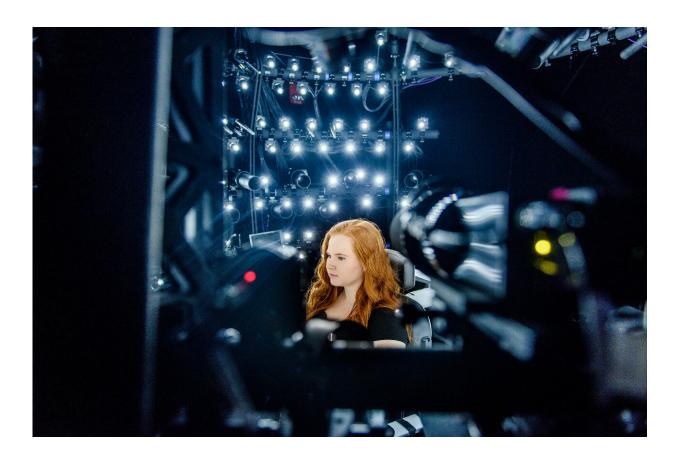# Virtual reality faces—animating precise, lifelike avatars for VR in real-time

June 25 2019

Creating lifelike avatars currently requires capturing large quantities of high-quality audio and video of an individual. Autumn Trimble is scanned in a highly customized system of cameras and microphones in Facebook Reality Lab's Pittsburgh office. Credit: Facebook

Computer scientists are focused on adding enhanced functionality to

make the "reality" in virtual reality (VR) environments highly believable. A key aspect of VR is to enable remote social interactions and the possibility of making it more immersive than any prior telecommunication media. Researchers from Facebook Reality Labs (FRL) have developed a revolutionary system called Codec Avatars that gives VR users the ability to interact with others while representing themselves with lifelike avatars precisely animated in real-time. The researchers aim to build the future of connection within virtual reality, and eventually, augmented reality by delivering the most socially engaged experience possible for users in the VR world.

To date, highly photo-realistic avatars rendered in real-time have been achieved and used frequently in computer animation, whereby actors are equipped with sensors that are optimally placed to computationally capture geometric details of their faces and facial expressions. This sensor technology, however, is not compatible with existing VR headset designs or platforms, and typical VR headsets obstruct different parts of the face so that complete facial capture technology is difficult. Therefore, these systems are more suitable for one-way performances rather than two-way interactions where two or more people are all wearing VR headsets.

"Our work demonstrates that it is possible to precisely animate photorealistic avatars from cameras closely mounted on a VR headset," says lead author Shih-En Wei, research scientist at Facebook. Wei and collaborators have configured a headset with minimum sensors for facial capture, and their system enables two-way, authentic social interaction in VR.

Wei and his colleagues from Facebook will demonstrate their VR real-time facial animation system at SIGGRAPH 2019, held 28 July-1 August in Los Angeles. This annual gathering showcases the world's leading professionals, academics, and creative minds at the forefront of

computer graphics and interactive techniques.

In this work, the researchers present a system that can animate avatar heads with highly detailed personal likeness by precisely tracking users' real-time facial expressions using a minimum set of headset-mounted cameras (HMC). They address two key challenges: difficult camera views on the HMC and the large appearance differences between images captured from the headset cameras and renderings of the person's lifelike avatar.

The team developed a "training" headset prototype, which not only has cameras on the regular tracking headset for real-time animation, but is additionally equipped with cameras at more accommodating positions for ideal face-tracking. The researchers present an artificial intelligence technique based on Generative Adversarial Networks (GANs) that performs consistent multi-view image style translation to automatically convert HMC infrared images to images that look like a rendered avatar but with the same facial expression of the person.

"By comparing these converted images using every pixel—not just sparse facial features—and the renderings of the 3-D avatar," notes Wei, "we can precisely map between the images from tracking headset and the status of the 3-D avatar through differentiable rendering. After the mapping is established, we train a neural network to predict face parameter from a minimal set of camera images in real time."

They demonstrated a variety of examples in this work, and were able to show that their method can find high-quality mappings even for subtle facial expressions on the upper face-an area that is very difficult to capture—where the camera angle from the headset is askew and too close to the subject. The researchers also show extremely detailed facial capture, including subtle differences in tongues, teeth, and eyes, where the avatar does not have detailed geometry.

In addition to animating the avatars in VR, the FRL team is also building systems that may one day enable people to quickly and easily create their avatars from just a few images or videos. While today's Codec Avatars are created automatically, the process requires a large system of cameras and microphones to capture the individual. FRL also aims to create and animate full bodies for expressing more complete social signals. While this technology is years away from reaching consumer headsets, the research group is already working through possible solutions to keep avatar data safe and ensure avatars can only be accessed by the people they represent.

 More information: "VR Facial Animation via Multiview Image Translation" SIGGRAPH 2019.

Provided by Association for Computing Machinery