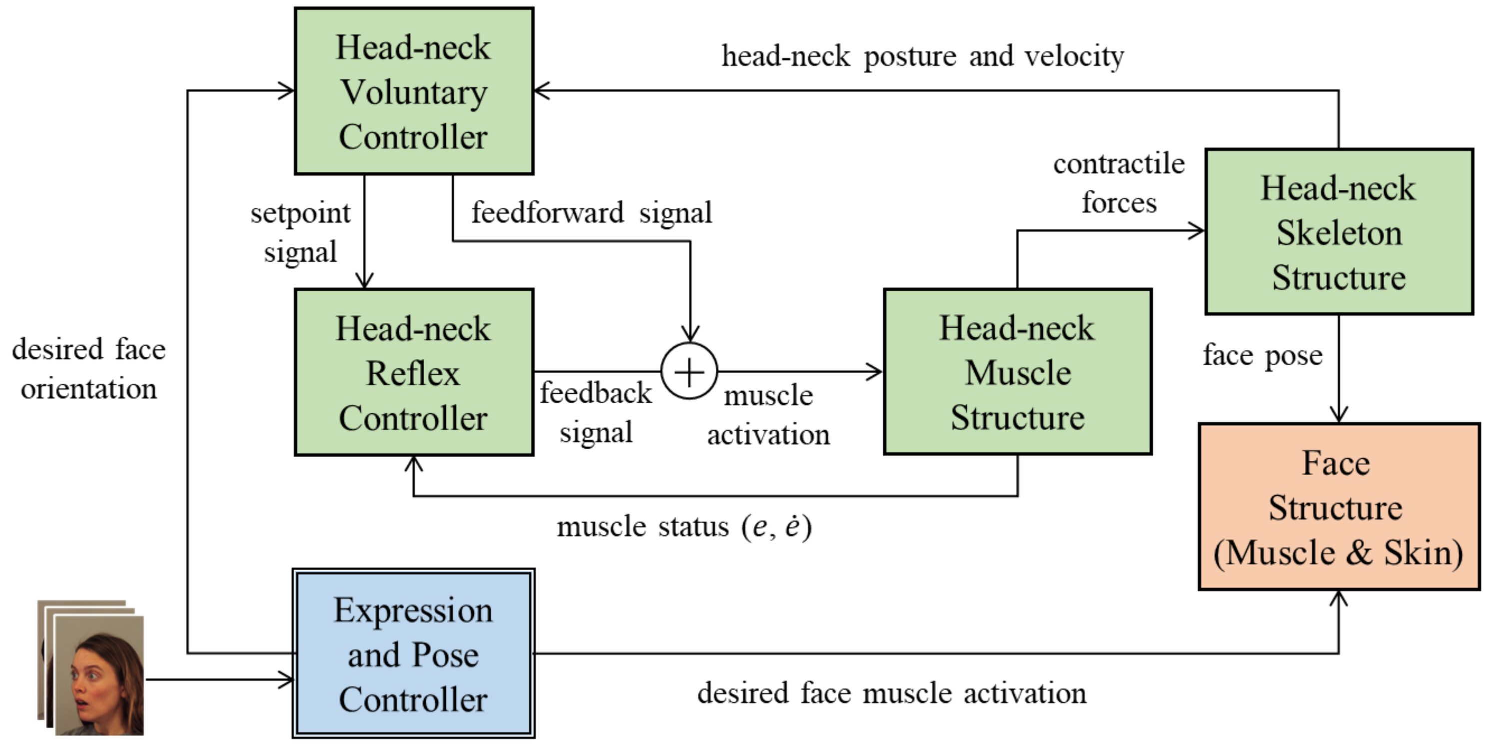
Facial Expression Transfer from Video Via Deep Learning

Xiao S. Zeng[1,*] (stevennz@cs.ucla.edu), Surya Dwarakanath[1,*], Wuyue Lu[1], Masaki Nakada[2], Demetri Terzopoulos[1]

[1] University of California, Los Angeles, [2] NeuralX, Inc

Overview

- We introduce the first biomechanical face-head-neck animation system that is capable of learning to reproduce expressions and head orientations through neuromuscular control.
- Our novel deep neuromuscular motor controller learns to map between FACS Action Units (AUs)[1] extracted from human facial images and videos and the activations of the muscle actuators that drive the biomechanical system.
- As a proof of concept, we demonstrate an automated processing pipeline for animating expressions and head poses using an improved version of the physics-based face-head-neck animation system developed by Lee and Terzopoulos[2], but which can potentially be applied to any physics-based, muscle-driven model.



Musculoskeletal Model

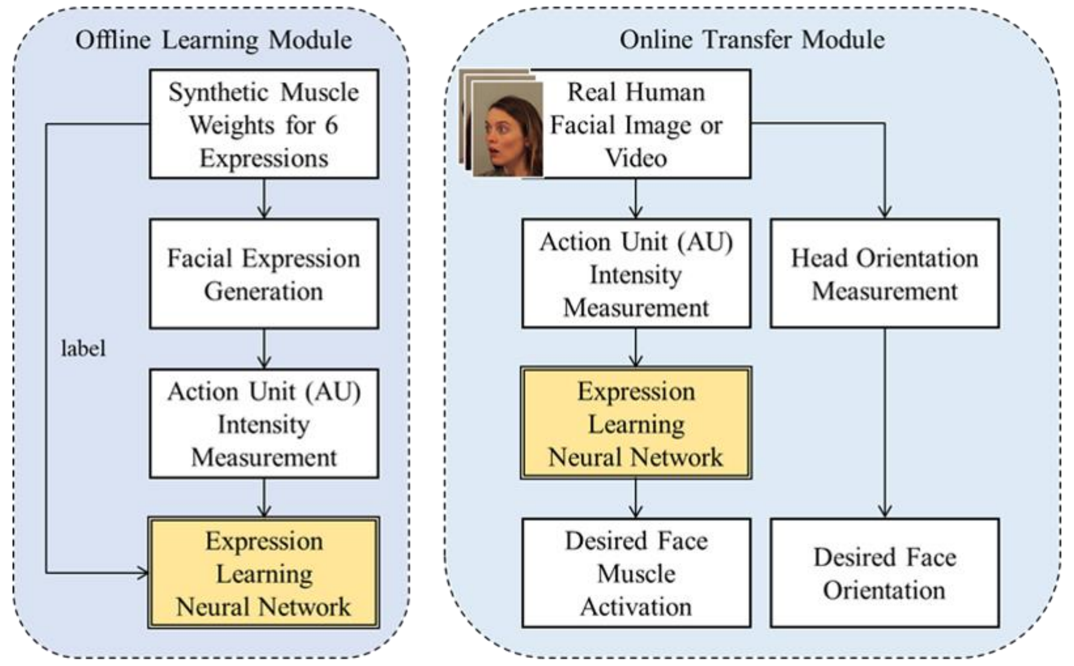
Our real-time musculoskeletal model is based on the work of Lee and Terzopoulos[2], but both the underlying face-head-neck control system and the facial expression system are significantly improved.

The skeletal structure is an articulated multibody dynamics system, with bones and joints consistent with human anatomy. The skeletal model is driven by a Hill-type muscle actuator model. The biomechanical face component consists of a facial soft tissue mode comprising synthetic skin and muscle layers together with a skull beneath them, all of which are constructed based on the work by Lee et al.[3]. The contractions of the embedded 26 pairs of muscles apply forces to the facial tissue layers, which deform to produce meaningful facial expressions. We augment the expressive details such as wrinkles on the face model by applying multiple levels of subdivision to increase the number of facial nodes that can be influenced by muscle forces. We also adapt a high resolution texture image to our generic face mesh to provide a more natural look.

Controlling The Model

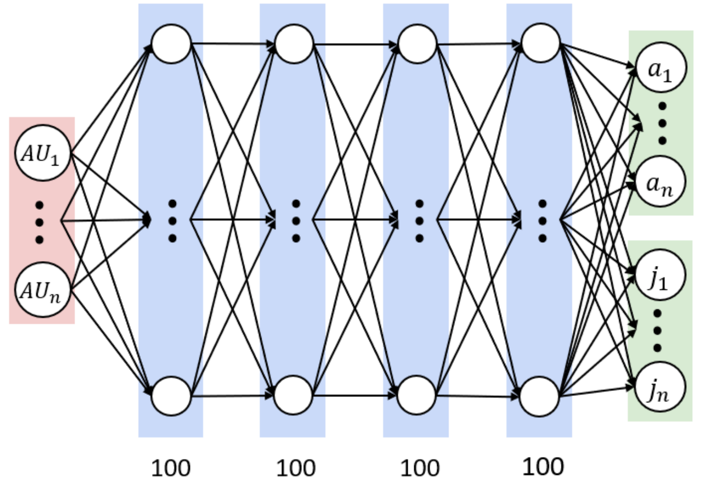
Offline Learning Module:

- The biomechanical face model generates expressions using synthetic muscle weights. Expression snapshots are processed by OpenFace toolkit to obtain AU intensities.
- AUs and synthetic muscle weights are used to train a DNN (yellow).



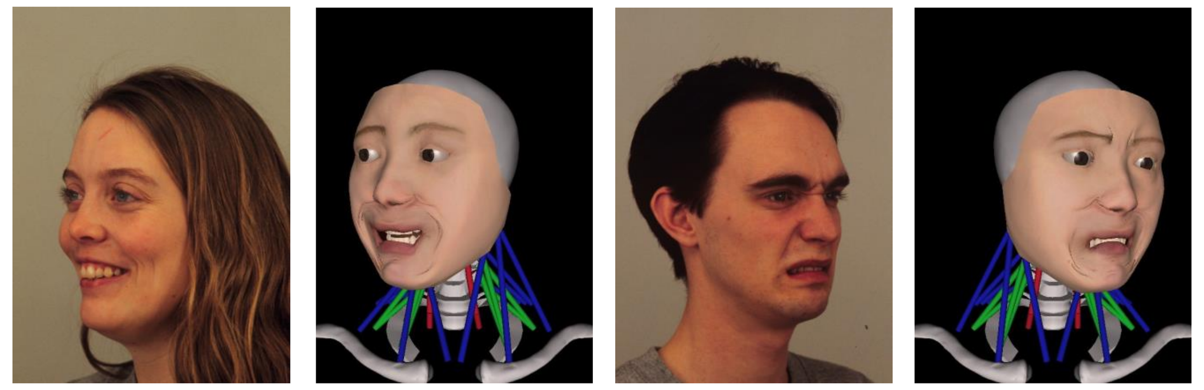
Online Transfer Module:

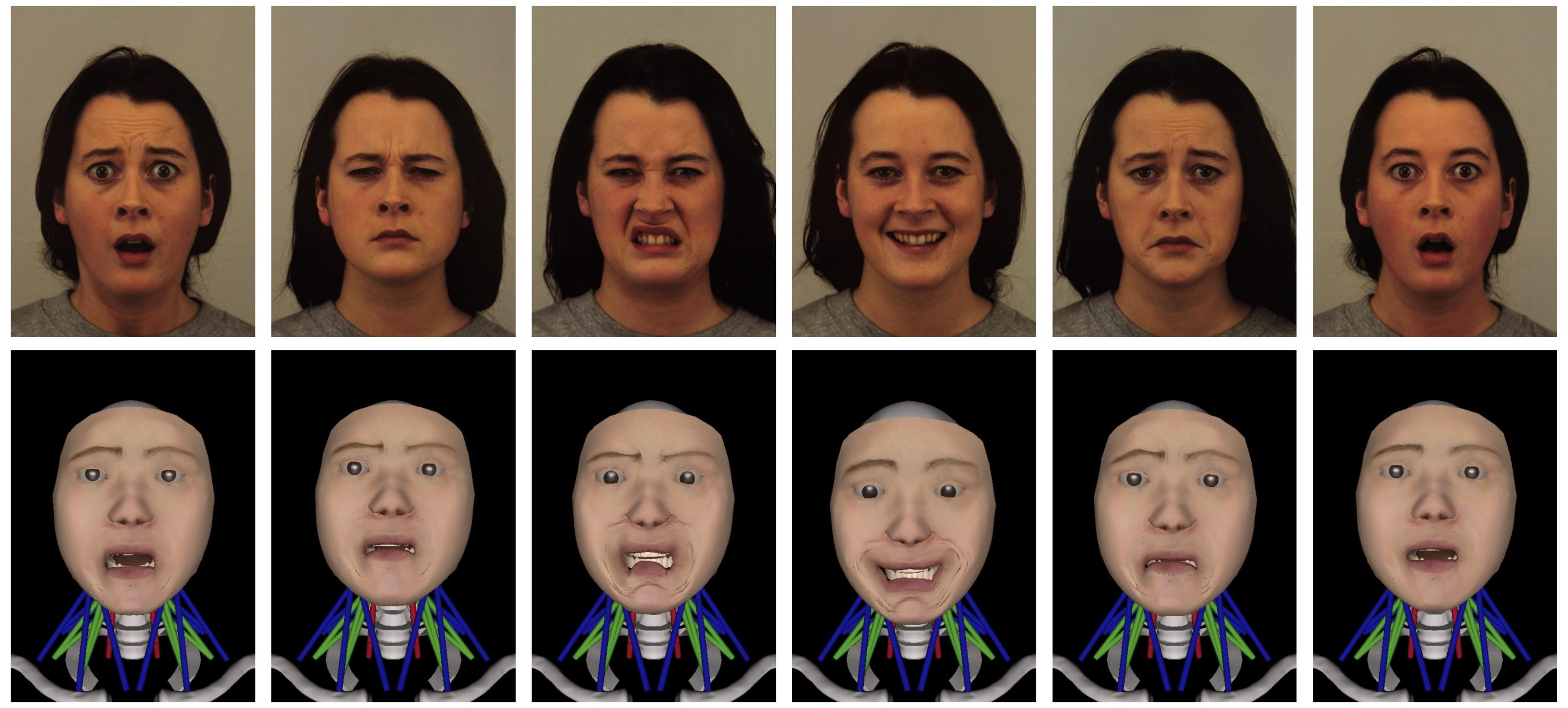
- OpenFace processes the input video to compute AU intensities and head pose.
- AU intensities are fed to pre-trained DNN to produce desired muscle weights. Head orientation is sent to head-neck system to produce corresponding head pose.



Results

We evaluate our expression and head pose transfer pipeline on different expressions and head orientations while using a variation of AUs and muscles in the biomechanical face-head-neck model. The figure below shows example transfer results of a male subject and a female subject enacting joy and disgust expressions with head orientation, and a female object enacting 6 basic expressions (fear, anger, disgust, joy, sadness and surprise) from the Karolinska Directed Emotional Faces dataset[4].





References

1. Jeffrey F Cohn, Zara Ambadar, and Paul Ekman. 2007. Observer-based measurement of facial expression with the Facial Action Coding System. The Handbook of Emotion Elicitation and Assessment 1, 3 (2007), 203–221.
2. Sung-Hee Lee and Demetri Terzopoulos. 2006. Heads up! Biomechanical modeling and neuromuscular control of the neck. In ACM SIGGRAPH 2006 Papers. 1188–1198.
3. Yuencheng Lee, Demetri Terzopoulos, and Keith Waters. 1995. Realistic modeling for facial animation. In Proceedings of the 22nd annual conference on Computer graphics and interactive techniques. 55–62.
4. Manuel G. Calvo and Daniel Lundqvist. 2008. Facial expressions of emotion (KDEF): Identification under different display-duration conditions. Behavior Research Methods 40, 1 (01 Feb 2008), 109–115.