Speech to Talking Face Animation
with Generative Adversarial Networks

Allen Tu, Jacob Ginzburg, Raymond H. Tu

Project Overview

Face Image

Speech Clip

Face Video Synthesizer

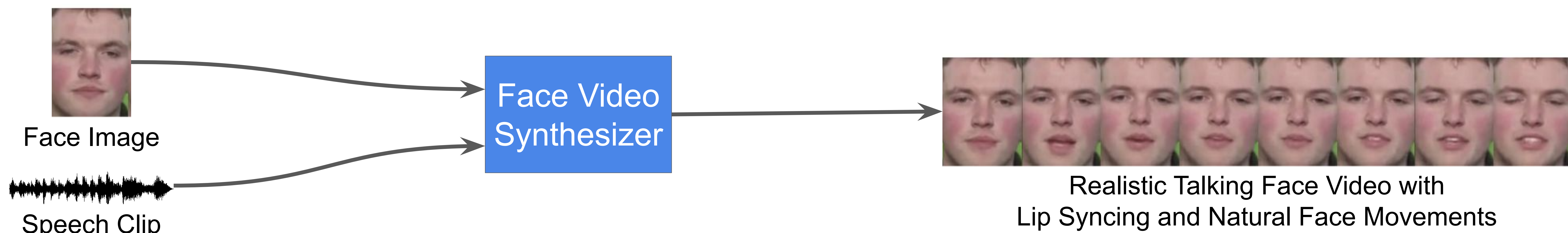Realistic Talking Face Video with Lip Syncing and Natural Face Movements
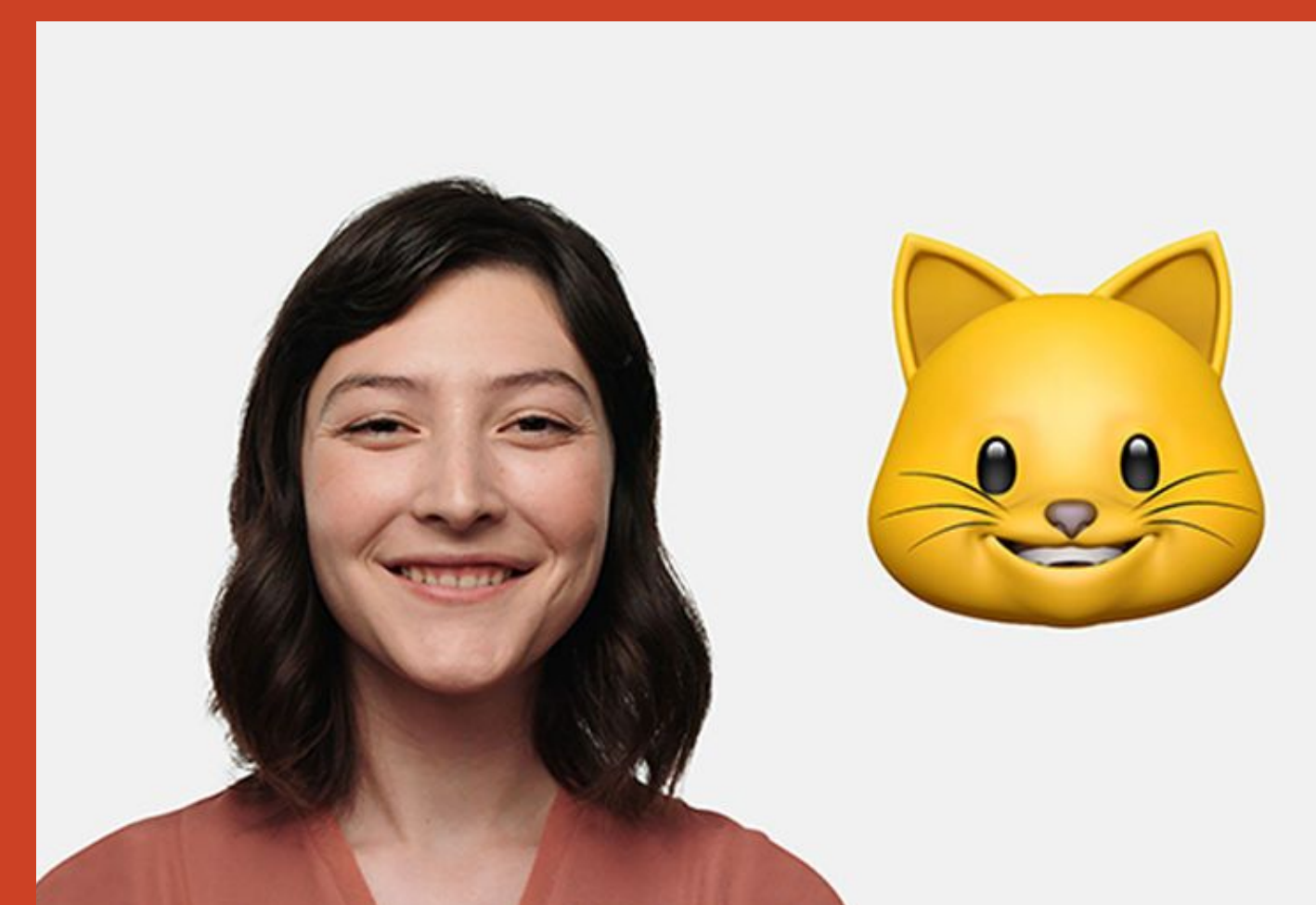
Applications

Animate computer generated character models
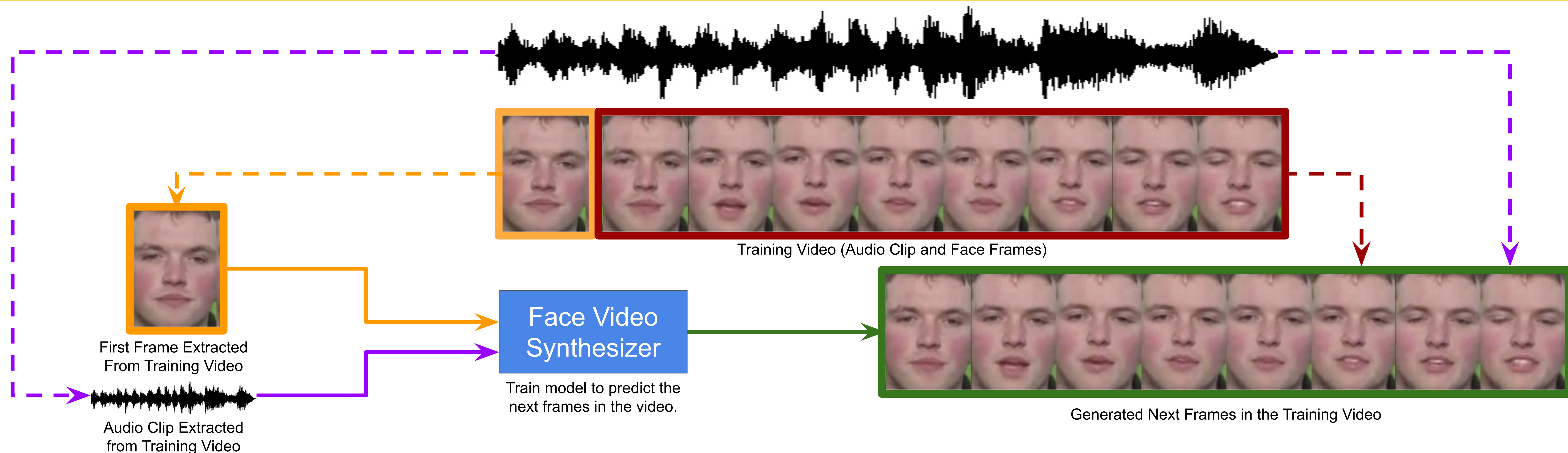
Reanimate lip movements in dubbed live action films

Reanimate lip movements in dubbed anime

Animate animojis, talking heads, and face filters

Generator Model

Training Video (Audio Clip and Face Frames)

First Frame Extracted From Training Video

Audio Clip Extracted from Training Video

Face Video Synthesizer

Train model to predict the next frames in the video.

Generated Next Frames in the Training Video

Discriminator Models

Model Generated Frames

Next Frames from Training Video

Audio Clip from Training Video

Frame Discriminator

Trained to determine if a frame is real.

Trains the Generator to generate realistic face frames.

Sequence Discriminator

Trained to determine if the sequence of frames in the video is real.

Trains the Generator to generate fluid and realistic videos.

Audio Sync Discriminator

Trained to determine if the lip movements in the video are synced to the audio.

Trains the Generator to generate face videos with lip and face movements that are synced to the audio.