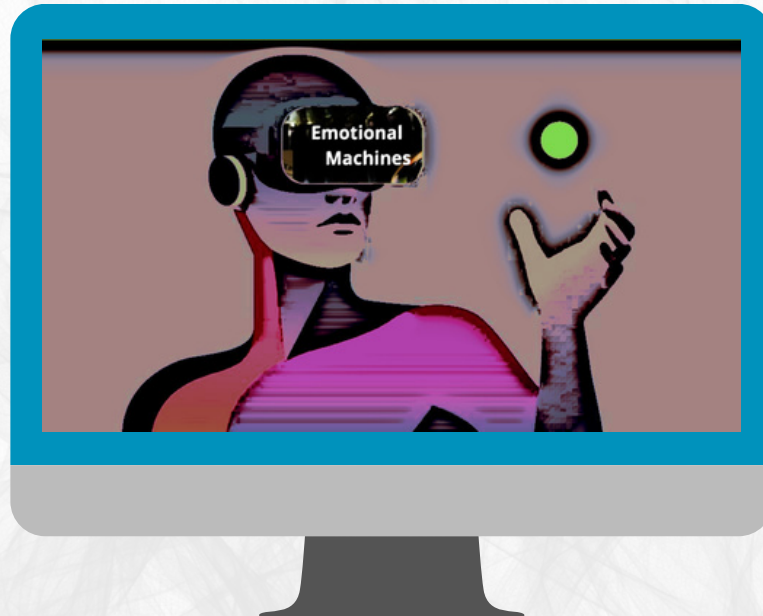


Emotional Machines

Multimedia Communications Technologies - MCT



Emotional Machines is an interactive media project that promotes the experience of affective virtual environments created through speech emotion recognition. In response to the existing limitations on emotion recognition models adopting computer vision, whose sources are usually hidden by a head-mounted display, we propose the adoption of speech as the user interface input. In particular, our system uses two machine learning models to predict three main emotional categories from high-level semantic analysis and low-level acoustics speech features. Emotions predicted are mapped to audiovisual representations by an end-to-end process that encodes emotions in virtual environments. We adopt a generative model of chord progressions to transfer speech emotion into music based on the tonal interval space. Images are built using a generative adversarial text-to-image synthesis. The generated image is then used as the style image in the style-transfer process onto an equirectangular projection of a spherical panorama selected for each emotional category. The result is an immersive virtual space encapsulating emotions in spheres disposed into a 3D environment. Thus, users can create new affective representations or interact with other previously encoded instances using the joysticks.