

Project Proposal

Project Title: Enhancing Emotional Expressiveness in Facial Animation from Speech

Company Advisor: Jordi Sanchez Riera

Institution: Institut de Robòtica i Informàtica Industrial (IRI), CSIC-UPC

Contact: jsanchez@iri.upc.edu

Project Overview:

The proposed research aims to improve facial animation techniques by enhancing the emotional expressiveness of digital avatars generated only from speech. Utilizing the Facial Landmark Animation and Modeling with Emotions (FLAME) as the foundational model, this project seeks to integrate advanced deep learning (DL) methods to disentangle pose and expression information, thereby achieving more nuanced and realistic animations.

Objectives:

- To refine the existing methods to better capture and render emotional subtleties in facial animations derived from speech data. Leveraging the FLAME model, we will introduce modifications to separate pose and expression data more distinctly, enhancing the model's ability to depict a wider range of emotional intensities.
- To implement cutting-edge audio encoders and integrate improvements from existing literature to enhance the synchronization between spoken emotions and facial expressions.
- To validate the improved model's effectiveness using Emotional Vertex Error (EVE) for emotion fidelity and traditional Lip synchronization error metrics.

Data Sources:

- HDTF and RAVDESS Datasets: These public datasets, comprising diverse video recordings with rich emotional speech content, will serve as the primary sources for training and testing our models.
- EMOCA Model: This will be employed to generate ground truth (GT) motion data against which our model's output will be benchmarked.

Evaluation Metrics:

- Emotional Vertex Error (EVE): This metric will assess the emotional accuracy of the facial expressions animated by our model.
- Lip synchronization Error: Traditional metric to evaluate the accuracy of lip movement synchronization with spoken audio.
- Additional Metrics: Exploration of other relevant metrics to further validate the expressiveness and accuracy of the animated faces.

Expected Outcomes:

The project aims to achieve significantly improved emotional expressiveness in facial animations, contributing to more lifelike and engaging digital interactions in applications ranging from virtual reality to digital communications and entertainment.

Impact:

Enhanced emotional expressiveness in digital avatars will not only improve user experience in interactive applications but also enhance the effectiveness of communication in virtual environments, providing a more intuitive and human-like interaction.