D6.4 INTERACTIVE FACIAL ANIMATION DEMONSTRATION

Grant Agreement nr	856879
Project acronym	PRESENT
Project start date (duration)	September 1st 2019 (36 months)
Document due:	28/02/2022
Actual delivery date	01/03/2022
Leader	Cubic Motion
Reply to	steve.caulkin@cubicmotion.com
Document status	Submission Version

Project funded by H2020 from the European Commission

Project ref. no.	856879
Project acronym	PRESENT
Project full title	**P**hotoreal **RE**altime **S**entient **ENT**ity
Document name	Interactive Facial Animation Demonstration
Security (distribution level)	Public (PU)
Contractual date of delivery	28/02/2022
Actual date of delivery	01/03/2022
Deliverable name	D6.4 Interactive Facial Animation Demonstration
Type	Demonstration
Status & version	Submission Version
Number of pages	7
WP / Task responsible	Cubic Motion (CM)
Other contributors	-
Author(s)	Ahmed Al-Obaidi, Emelia Fiell, Jack Saunders, Francisco Peñaranda and Steve Caulkin
EC Project Officer	Ms. Diana MJASCHKOVA-PASCUAL Diana.MJASCHKOVA-PASCUAL@ec.europa.eu
Abstract	This Report describes the research carried out to generate facial animation for the digital agent. Our methods, progress and results are documented.
Keywords	Facial Animation, Digital Human
Sent to peer reviewer	Yes
Peer review completed	Yes
Circulated to partners	No
Read by partners	No
Mgt. Board approval	No

Document History

Version and date	Reason for Change
1.0 06-01-2022	Document created by Emelia Fiell
2.0 18-02-2022	Submitted for peer review
3.0 28-02-2022	Final Version

Table of Contents

1 EXECUTIVE SUMMARY

The work demonstrated in this deliverable covers the generation of facial animation for the digital agent. An intermediate agent asset was built, which was used during the development process whilst Framestore built the high resolution agent. This also provides an alternative, less computationally expensive agent for use within the wider project. A facial animation pipeline was developed to process training data captured for animation generation research. Systems were then built to create animation for the speech and emotion from higher level input cues from other modules in the reference implementation. These are presented as video demonstrations in the Unreal Engine project reference implementation and UE5 using MetaHuman characters.

You can find the video demonstration here:

https://epicgames.box.com/s/u17t1k6zf95gds974lm9wj082ib4mmgq

2 BACKGROUND

This deliverable is a demonstration of the work described mainly in D3.5 Visual Facial Animation Report, focusing on the interaction with the wider system in Unreal Engine. The methods described here will be refined and applied in WP8T4 Prototype Evaluations and WP9T3 Agent Demonstration.

3 FACIAL ANIMATION DEMONSTRATION

The main objective of this deliverable is to show the facial animation system operating in context and show how the motion generation component interacts with other modules. The demonstrations here are shown in the Unreal Engine 4.27 project reference implementation as well as Unreal Engine 5. Since UE5 is likely to become the industry standard, we decided to develop for both platforms.

In order to ensure that we are able to support a diverse range of characters, the animation systems were built around the Unreal Engine MetaHuman control rig. This means that the animation produced is natively compatible with all MetaHumans, as shown in some of the UE5 examples.

The demonstrations are presented in the attached video and are described below:

3.1. Intermediate Agent

This clip shows the intermediate agent in the UE4.27 reference implementation for the project. The animation here is from the broadcast use case.

The intermediate agent has been developed as a lightweight alternative to the high resolution agent produced by Framestore. The lightweight body rig was produced using the newly released in-house tools of the UE MetaHuman technology. MetaHuman is a tool created by Epic Games that will empower the user to create a bespoke photorealistic digital human, fully rigged and complete with hair and clothing. The switch to using MetaHuman technology enables the creation of a generic skeleton and a corresponding mesh using auto rigging methods which empower PRESENT with several advantages. The benefits of MetaHuman include greater character diversity, cross-platform adaptability, wider applications and use cases and the possibility of continuous expansion as the tool advances.

3.2. Facial Geometry Solve

This clip shows the facial geometry solve for an example shot. This involves first tracking the 2D image contours representing key facial features, then reconstructing the depth map of the face from the stereo cameras, and finally fitting a coherent 3D model of the facial geometry to the sequence. This is intended to give an accurate representation of the facial performance, from which we can estimate the character animation curves.

3.3. BML Interface

This clip shows the implementation of the BML (behaviour markup language) interface to pass instructions to the agent.

A separate BML Realiser Component was created in the reference implementation to handle BML parsing. This avoids having to implement a parser in the Motion Generation Component and separates the logic required to deal with arbitrary multiple cues coming from the wider system from the motion generation itself. The BML Realiser Component is then able to directly call input methods of the Motion Generation Component.

3.4. Speech

This clip shows an example solved speech sequence from the broadcast use case. The fitted mesh from the geometry solve step has in turn been solved to the character rig controls. This is done using a method designed to produce semantically meaningful animation curves based on training data provided by an animator. This approach automatically produces high quality animation data, which can be used for speech sequences or as training data for generative models.

3.5. MetaHuman Speech

This clip shows another example speech performance from the broadcast use case applied to two MetaHuman characters in UE5. This shows that the animation is transferable to MetaHumans and other compatible rigs with minimal loss of quality. The audio here is taken from the male actor's performance; we have also captured female audio data to increase the diversity of agents the system can support.

3.6. Emotional Model

This clip shows the emotional model driving the character (in this case a MetaHuman) in UE5. The emotional valence and arousal values are connected to a state model, which uses training animation data from the motion capture sessions designed to represent specific emotions. This approach allows the emotion of the agent to be driven by cues from the wider system.

3.7. Style Transfer

In order to further improve the emotional model and investigate the effect of identity on facial animation, research work has been carried out into using a Generative Adversarial Network (GAN) to transfer animation style. This model is trained to transfer animation data between different emotional states (neutral, happy and sad) and between two actors (Gareth and Fran). The clip shows this model applied to an input animation, which is transferred to a MetaHuman in UE5 using the emotional and identity parameters to manipulate the animation to the desired style.

4 CONCLUSION

This deliverable shows that we have successfully implemented a facial animation system capable of driving the digital agent. This is based on an automated facial animation pipeline, which was used to build a high-quality training data set of facial animation.

The integration work with the UE reference implementation led to the creation of a flexible BML realiser component, which handles the passing of high level cues from other parts of the system to the motion generation component. Each of the main animation systems has been integrated with the motion generation component. Work will continue on these components to support project use cases.

Further work has been done to integrate the facial animation systems with UE5. This will be beneficial to the exploitation of the project as UE5 becomes more widely used and will allow the system to use some of the improved capabilities in the engine itself. The decision to base the whole animation system on the MetaHuman control rig significantly expands the scope to use the project for anyone who has access to those characters and allows us to drive a much more diverse range of characters.

5 ACRONYMS AND ABBREVIATIONS

Mo-Cap	Motion Capture
UE	Unreal Engine
FK	Forward Kinematics
IK	Inverse Kinematics