D3.2 INTERIM VISUAL FACIAL ANIMATION DEMONSTRATION

Grant Agreement nr	856879
Project acronym	PRESENT
Project start date (duration)	January 1st 2018 (36 months)
Document due:	30/11/2020
Actual delivery date	30/11/2020
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	Interim Visual Facial Animation Demonstration
Security (distribution level)	Public
Contractual date of delivery	30/11/2020
Actual date of delivery	30/11/2020
Deliverable name	D3.2 Interim Visual Facial Animation Demonstration
Type	Demonstration
Status & version	Submission Version, v1.2
Number of pages	8
WP / Task responsible	Cubic Motion
Other contributors	
Author(s)	Francisco Peñaranda, Steve Caulkin, Ahmed Al-Obaidi, Emelia Fiell
EC Project Officer	Ms. Adelina Cornelia DINU - Adelina-Cornelia.DINU@ec.europa.eu
Abstract	This Deliverable describes the full pipeline developed to produce facial animation on the virtual agent from the performance of the real actor. The key innovations introduced in the different elements of the workflow are presented together with encouraging preliminary results.
Keywords	facial animation, performance capture system, 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 13-11-2020	Document created by Francisco Peñaranda
1.1 16-11-2020	Version for internal review (14 days before submission date)
1.2 30-11-2020	Revisions in response to review: final version submitted to Commission

Table of Contents

1 EXECUTIVE SUMMARY

The main elements of the pipeline developed by Cubic Motion to create facial animation from the actor's captured performance are presented in this Deliverable. They include some key innovations that have been developed up to a good state: a new stereo head-mounted camera capture system, the creation of a lightweight asset or facial rig of the actor Gareth, and a new approach to create the facial animation from the actor's performance.

The first version of the full pipeline has been implemented and tested on a relatively small dataset of captured footage, limited by the COVID-19 restrictions. The preliminary results are very encouraging and demonstrate the great potential of the established approach, but it will require more improvements when more data from the actor can be recorded. This workflow will be used to create a reference database of realistic facial performance from the actor Gareth in a wide range of acting situations. Some elements of the pipeline and the created facial animation database will be integrated with the rest of modules to drive the face of the virtual agent in the first complete versions of the engine, this will be the main focus over the coming months.

2 BACKGROUND

D3.2 Interim Visual Facial Animation Demonstration is the second Deliverable associated with the work package *WP3 Agent Visual Creation* and the first one related to the task *WP3T3 Facial Animation for Virtual Agents*. It shares some important elements of the virtual agent developed by Framestore inside the WP3, but it represents a fairly independent unit that incorporates key novel functionalities in the process towards driving the digital avatar in a realistic fashion.

The main aim of this Deliverable is to present the basic building blocks that have been developed to capture the actor's performance and create a realistic virtual version of his facial expressions. In addition, it shows the first results obtained, which demonstrate the great potential and the relatively good state of the facial animation pipeline.

The work presented here will be extended in M30 in the Deliverable *D3.5 Visual Facial Animation Report*. Moreover, it will have special relevance for the rest of tasks inside the WP3 and other work packages, like *WP6T3 Interactive Facial Animation* and *WP8T4 Prototype Evaluations*.

3 INTRODUCTION

This Deliverable describes the main elements developed and used to create the underlying mechanisms that will animate the face of the virtual agent. The characteristics of these elements are detailed in section 4. Two videos have been created to accompany and illustrate this Deliverable, as will be properly appointed in section 4. The links to the accompanying videos can be found in the section 7 (Annex).

3.1 Main objectives and goals

- Establish a pipeline to create facial animation of the virtual agent from the offline captured performance.
- Customise this pipeline to provide the most optimal results for the actor Gareth for a wide variety of facial expressions and acting situations.
- Create a lightweight digital asset of the actor Gareth to improve the efficiency of the pipeline.
- Create a reliable and comfortable capture system to record high-quality stereo footage from the actor's facial performance.
- Implement new algorithms that make use of the extracted geometrical and temporal information to solve the control values of the lightweight asset.

4 FACIAL ANIMATION

The animation of the face of the virtual agent is a relatively complex process that can be decomposed or simplified into a few high-level elements. Figure 1 outlines these components and their relationship inside the whole workflow.

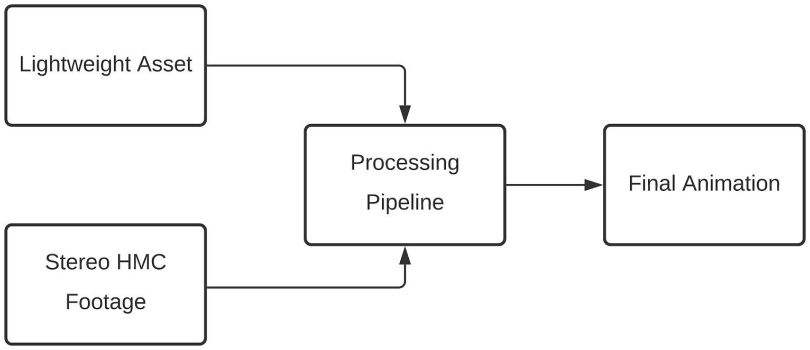

Figure 1. Facial animation flowchart

The main components and subcomponents that can be identified are:

1. **Lightweight asset**: it is one of the two main inputs of the facial animation process and is also referred to as *intermediate-level facial rig*. It is a virtual entity that provides a high-level abstraction of the face topology and tries to simulate plausible movements and deformations of the face. The so-called *rig logic* encodes several interrelated layers that define how the different regions of the face must move coherently to approximate the physics behind a finite range of facial expressions.

2. **Stereo HMC footage**: this is the other main input of the facial animation workflow. It consists of two synchronised videos of the actor's facial performance recorded with a new proprietary Head-Mounted Camera (HMC) capture system.

3. **Processing pipeline**: this includes all the algorithms that take the previous elements as inputs to produce the final facial animation. This module adjusts the facial topology of the rig to match the geometrical performance of the actor's face for each frame of the recorded footage. During this process, the specific control values of the rig logic are computed or solved.

4. **Final animation**: the main output of the processing pipeline is a set of animation curves that reflect the values of the controls of the lightweight asset for each frame of the captured facial performance.

Cubic Motion has taken advantage of the PRESENT Project to introduce several key novelties at different levels of this workflow with respect to previous facial animation approaches. The following sections summarise the main advancements and provide more details of each main element of the facial animation workflow.

4.1 Lightweight asset

The lightweight asset or intermediate-level facial rig is one of the central input pieces in the facial animation pipeline. Creating a rig with a lower level of detail than the *high-quality asset* developed by Framestore increases the performance of the whole facial animation process. In addition, having access and controlling the specific design of the rig logic allows us to leverage the full potential of the processing pipeline.

Cubic Motion used internal rigging tools to incorporate some specific data from the actor Gareth that was supplied by Framestore. In particular, Framestore provided a mesh of the actor Gareth in a neutral pose, which defines the reference geometry of his face, and an initial version of his skin texture to give the asset a more realistic look. Framestore also kindly supplied other data acquired to create the high-quality asset, like a range of high-quality 4D scans of Gareth. In this first version of the lightweight asset, only the neutral mesh and the skin texture were employed to create a customised version of an in-house generic facial rig.

Figure 2 shows two views of the rendered lightweight asset. For a better illustration of the rig functionalities, a video entitled D3.2_expressions.mp4 that accompanies this Deliverable shows the facial expressions that the lightweight asset can produce.

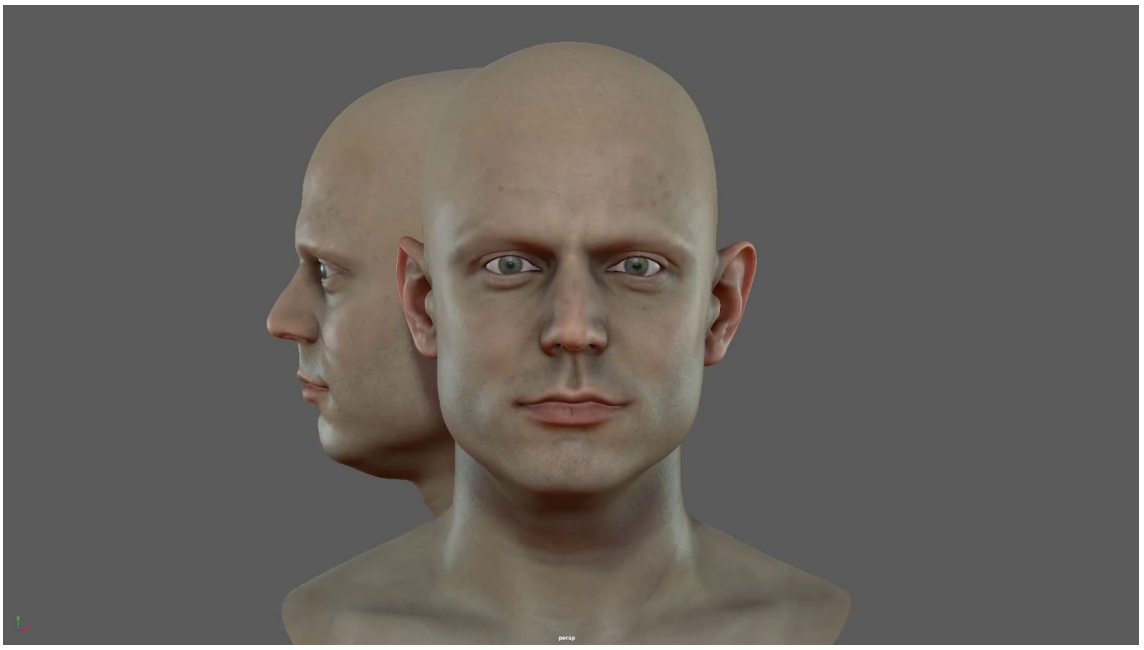

Figure 2. Side and front view of the facial lightweight asset

4.2 Stereo HMC footage

The second main input of the process to generate facial animation from the actor's performance is the recorded footage of the actor's face. An HMC capture system is normally used to register the performance without restricting the body movements of the actor too much. The PRESENT Project has given Cubic Motion the opportunity to make key progress on the design and testing of a new stereo HMC capture system.

The new prototype HMC includes two main novelties:

1. A new pair of front stereo cameras that provide synchronised high-quality images recorded at a high frame rate.

2. A new mechanical design of the camera rig and the helmet. This new HMC version is constructed with lighter composite materials that both reduce the overall weight and maintain good stability for the system.

A real prototype version of the new stereo HMC was used during the last capture session with the actor Gareth on 15 October 2020 at Cubic Motion's motion capture studio. The actor wore the HMC for more than 6 hours and his feedback was very positive.

4.3 Processing pipeline

The processing pipeline comprises all the core algorithms that take the information encoded in the lightweight asset and the stereo HMC footage as inputs to generate the final facial animation that will mimic the actor's performance. This pipeline represents a novel approach with respect to previous facial animation workflows at Cubic Motion, thanks to two new main factors. The first one is the use of stereo cameras, which allows to extract more detailed geometric information from the actual performances. The second one is to have access to the underlying rig logic of the lightweight asset, which impacts on how the rig control values are solved attending to objective geometric optimisations. The main benefit of this innovative approach is that it does not require any supervision from a skilled Animator, which is a costly and scarce resource.

4.4 Final animation

The final result of the processing pipeline is a set of values for the controls that compose the rig logic of the lightweight asset for each frame. These values can be imported into the virtual rig to create the final animation.

Figure 3 shows an example frame of a rendered scene that includes a reference camera view at the left and two views of the facial rig at the right. The complete example animation is included in the video, D3.2_animation.mp4.

Figure 3. Example frame of the final animation

5 CONCLUSION

This Deliverable describes the full pipeline that has been developed at Cubic Motion to create facial animation of the virtual agent from the recorded actor's performance. The main tasks during the first 15 months of the Project have been focused on introducing novel elements with respect to previous facial animation pipelines at Cubic Motion. These innovations include a completely new stereo HMC capture system, a lightweight asset of the actor, and the use of a completely new approach to solve the control values of the facial rig, which does not require the supervision of skilled Animators.

Due to the COVID-19 restrictions, only two capture sessions have been carried out with the actor Gareth so far and the last one occurred less than a month before writing this Deliverable.

Therefore, the data that was available to create and test all the elements in the pipeline was very limited. Nevertheless, as can be seen in the accompanying videos, the preliminary results obtained with this workflow show overall high levels of fidelity with respect to the real performance of the actor. These results are very encouraging and demonstrate the huge potential of the defined facial animation pipeline.

The next steps will be mainly focused on testing this pipeline with more data captured from the actor Gareth and improving its different key elements. During this iterative process of refinement, a database of a wide variety of facial animations from the actor's performance will be created. This database will be incorporated into the reference implementation to drive the face of the virtual agent in a realistic way in the different use case scenarios.

Another key step will be to integrate the lightweight asset into the reference implementation developed by Framestore in Unreal Engine and create the modules that will generate different facial performances in that environment. Finally, Cubic Motion will develop a methodology to map the controls of the lightweight asset to the high-quality asset developed by Framestore in order to drive it in a similar fashion as the lightweight asset.

6 ACRONYMS AND ABBREVIATIONS

HMC: Head-Mounted Camera

7 ANNEX

Links to the accompanying videos:

- D3.2_animation.mp4: https://drive.google.com/file/d/1gWbYAvLSQrfEI3srLSNrEECl56cDG8eE
- D3.2_expressions.mp4: https://drive.google.com/file/d/1ZxA_BoKq1UBQpvo2OEGICqFpGZEorY54