Sign Language Translation Mobile Application and Open Communications Framework

Deliverable 5.1: A Virtual Character

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 101017255

Project Information
Project Number: 101017255
Project Title: SignON: Sign Language Translation Mobile Application and Open Communications Framework
Funding Scheme: H2020 ICT-57-2020
Project Start Date: January 1st 2021

Deliverable Information
Title: A Virtual Character
Work Package: WP 5 - Target Message Synthesis
Lead beneficiary: UPF
Due Date: 30/06/2022
Revision Number: V0.1
Authors: Josep Blat, Víctor Ubieto
Dissemination Level: Public
Deliverable Type: Demonstrator

Overview: In this document the 3D character that will be used in the project is presented. Virtual avatars are one of the most delicate aspects in sign synthesis, as the skeleton and blendshapes should be good enough to support high quality animations, and the visual quality should be very good, to get users' acceptance. We discuss the improvements over Eva (our previous avatar), mainly in a skeleton with more detail in hands and blendshapes supporting high quality expressivity and mouthing, compatible with our

sign synthesis pipeline discussed in D5.8, and much higher quality in the appearance, with special attention to skin, eyes and hair, with better overall textures in terms of Physically Based Rendering.

Revision History

Version #	Implemented by	Revision Date	Description of changes
<V0.1>	<Víctor Ubieto>	<14/06/2022>	Shell version
<V0.2>	<Víctor Ubieto>	<23/06/2022>	After reviewed by partners

The SignON project has received funding from the European Union's Horizon 2020 Programme under Grant Agreement No. 101017255. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the SignON project or the European Commission. The European Commission is not liable for any use that may be made of the information contained therein.

Approval Procedure

Version #	Deliverable Name	Approved by	Institution	Approval Date
V0.2	D5.1	Aoife Brady	DCU	24/06/2022
V0.1	D5.1	Andrea Cavallaro, Marcello Scipioni	FINCONS	17/06/2022
V0.1	D5.1	Vincent Vandeghinste	INT	20/06/2022
V0.1	D5.1	Olatz Perez de Viñaspre	UPV/EHU	21/06/2022
V0.1	D5.1	John O'Flaherty	MAC	17/06/2022
V0.1	D5.1	Euan McGill	UPF	22/06/2022
V0.1	D5.1	Irene Murtagh	TU Dublin	22/06/2022
V0.1	D5.1	Mathieu De Coster	UGent	21/06/2022
V0.1	D5.1	Jorn Rijckaert	VGTC	20/06/2022
V0.1	D5.1	Ellen Rushe	NUID UCD	20/06/2022
V0.1	D5.1	Henk van den Heuvel	RU	17/06/2022
V0.2	D5.1	Myriam Vermeerbergen	KU Leuven	27/06/2022
V0.1	D5.1	Davy Van Landuyt	EUD	20/06/2022
V0.1	D5.1	Mirella De Sisto	TiU	27/06/2022

Acronyms

The following table provides definitions for acronyms and terms relevant to this document.

Acronym	Definition
SL	Sign Language
MF	Manual Features
NMF	Non Manual Features
WFD	World Federation of the Deaf
WASLI	World Association of Sign Language Interpreters
BML	Behaviour Mark-up Language
PBR	Physical-Based Rendering
GPU	Graphics Processing Unit

Table of Contents

1. Introduction

The output sign language translation that SignON intends to provide is by means of the animation of a virtual/synthetic character who signs. A virtual character who signs is usually called a *Virtual Signer* or *Signing Avatar.* Both terms are used as synonyms throughout this report, although *Avatar* could more specifically denote the representation of the self by means of a virtual character. Taking into account this aspect of self representation or the adaptation to different usages, it is convenient that the signing character is *personalisable*, i.e., modifiable by the users designing the translation system implementation, or the users who are represented by the virtual character. This document is the first deliverable related to the design/modelling of the character or avatar, due in M18 of the project, and details the development of the basic avatar that will be used throughout the project. The personalisation of this basic character, modifying, for instance, gender, skin colour, hair, clothes, etc. will be the focus of the second deliverable, due in M36. The personalisation of the character will be *co-designed* with the end users, in line with the co-creation fundamental principle of the SignON project.

There are several *technical* aspects that need to be addressed in the design, which are related to the skeleton of the character. To perform computer animation, an articulated body (such as that of a human or an animal) is usually represented by means of a *hierarchical* structure of *segments* and *joints*, which is denoted as its *skeleton*. In this deliverable we discuss the skeleton requirements that make it usable for Sign Language (SL) animation, mainly related to a very detailed representation of the hands, as *Manual Features* (MFs) are most important in SLs along with the Non-Manual Features (NMFs). A second aspect to be considered is the importance of the facial expressions for the understanding of the messages to be conveyed; NMFs are also important in SLs. The modelling of the face should be appropriate to be able to support rich *facial expressivity* and suitable *mouthing*. On the other hand, *interoperability* with other systems is also very important, in the context of open systems and solutions of the SignON project.

On the other hand, the character's *appearance* will significantly impact interaction with users, and its design needs to be grounded on their requirements, which might be diverse for different users and use cases.

This report addresses the different issues indicated in the previous two paragraphs, as well as providing their implementations in the actual virtual character, the (basic) virtual avatar for the SignON project.

Section 2 provides the main background against which the design of the character has been undertaken. Section 3 is focused on the skeleton, while Section 4 on the aspects of appearance and rigging of the character, illustrated by some images of partial aspects and the initial character that was used.

The whole redesign of the character, with support of a specialist, is finished except for the final hair textures, which will be available in M19 and reported in the M36 deliverable. The character is an .OBJ[1], with the texture files, on which personalisation aspects will be developed.

[1] A very common file format for assets in Computer Graphics (https://en.wikipedia.org/wiki/Wavefront_.obj_file)

2. Background Concepts on Virtual Characters Design

In this section we discuss some concepts related to the context where the basic character is designed, and the corresponding challenges faced. A full description of the character's design is essentially contained in section 4.2.

2.1. The Uncanny Valley

The *Uncanny Valley* is a not completely precise concept related to the difficulty of creating characters which are realistic but do not look "creepy". Figures such as Figure 1 are used to depict the phenomenon[2]. The image represents that acceptance of a virtual character by users increases from cartoonish style to more realistic one, until a huge acceptance drop (a *valley*) happens when getting very close to real representation.

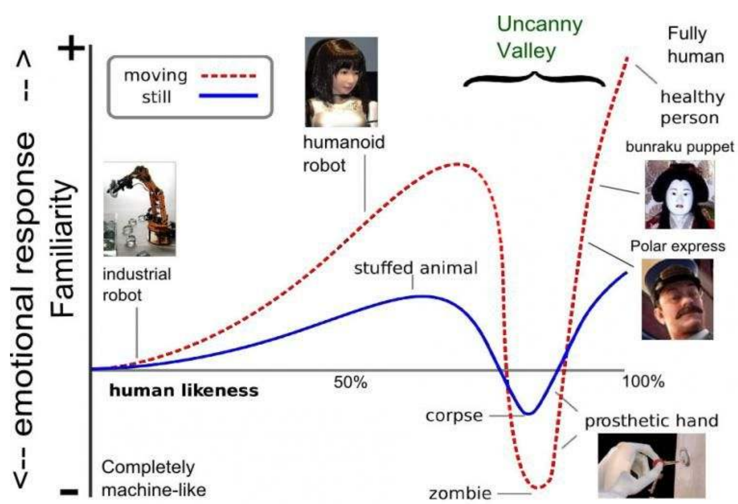

Fig 1: Representation of the Uncanny Valley effect with specific examples

There is diverse evidence on the Uncanny Valley[3], which itself can denote different aspects. But the most important factor driving the effect seems to be the mismatch(es) the user can perceive, that impact on

[2] Appearing in https://loperntu.github.io/ai2017/ which references http://www.xenosystems.net/uncanny-valley/

[3] Probably the most solid paper is a 2015 review: Kätsyri J, Förger K, Mäkäräinen M and Takala T (2015) A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. *Front. Psychol.* **6**:390. doi: 10.3389/fpsyg.2015.00390.

The more recent paper Mincheol Shin, Se Jung Kim, Frank Biocca, The uncanny valley: No need for any further judgments when an avatar looks eerie, *Computers in Human Behavior*, **94**, p. 100-109, 2019, https://doi.org/10.1016/j.chb.2019.01.016, does not necessarily provide more clarity to the issue.

how the users unconsciously respond. A virtual character with a cartoonish look, little intelligence, and lacking sophisticated animation might appear as more believable than one with a very high quality look, but not having the same level of reactivity or animations. Figure 2 shows another representation of the Uncanny Valley challenge, Realism vs Naturalism/Acceptance.

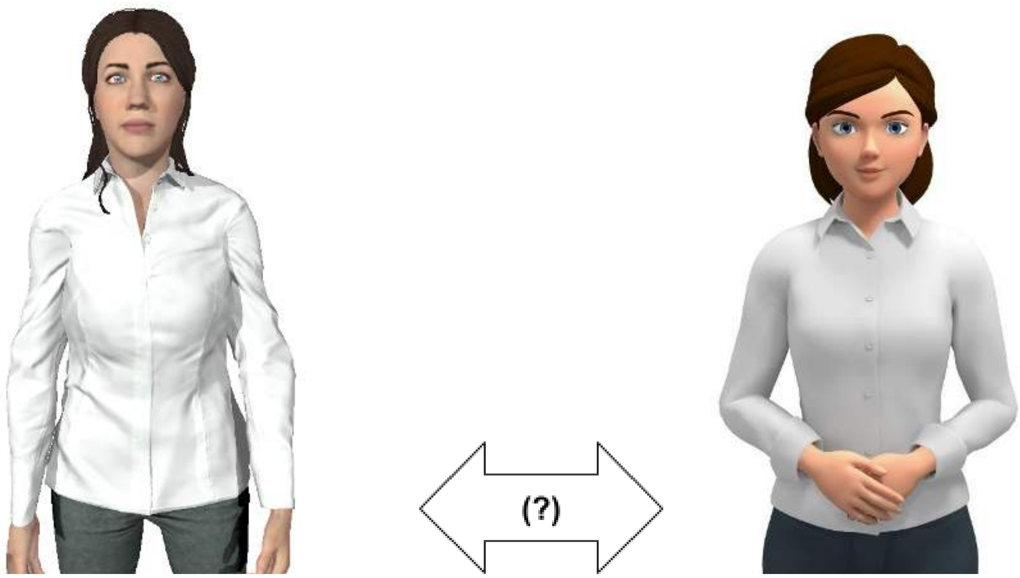

Fig 2: Comparison of different 3D avatars. **Left:** Current Eva. **Right:** Recreation of Eva for SiMAX Project

2.2. The Deaf / Hard of Hearing Users Requirements

Using virtual signers for translation is controversial for Deaf and Hard of Hearing (DHH) users, as illustrated by the joint declaration of WFD and WASLI[4] on signing avatars. The SignON project has been carrying out extensive work with users to elicit requirements that might make the virtual signer more acceptable to users, reflected, for instance, in D1.3 *First User Requirements Report*. A very recent paper by Quandt et al.[5] discusses different aspects of the perception of avatars by DHH users, which hint at the very nuanced different dimensions of the issue. There are two visual aspects to make the signing avatar more acceptable: better animations, as the current ones look far from being acceptable; and much higher, and customisable, visual quality. Other critical components of acceptance that have been identified, use-context and quality of translation, are more complex, and go beyond the reach of this

[4] https://wfdeaf.org/news/wfd-wasli-issue-statement-signing-avatars/

[5] Quandt LC, Willis A, Schwenk M, Weeks K and Ferster R (2022) Attitudes Toward Signing Avatars Vary Depending on Hearing Status, Age of Signed Language Acquisition, and Avatar Type. *Front. Psychol.* **13**:730917. doi: 10.3389/fpsyg.2022.730917

task, although customisation of the avatar and its background could support very partially use-context. The animation shortcomings probably stem from the fact that the bulk of the research and development of virtual characters has focused on visual aspects related to entertainment, and comparably very little in the fine grained animation required for SLs. Animation aspects are the central focus of another task of WP5, T2 *Developing an interactive system of learning from user generated signed content* and thus will not be discussed further in this report. Rendering quality is part of the visual one, it is the subject of WP5T5 *Real-time synthesis and delivery of target SL* and will not be discussed either. On the other hand, the avatar should be adaptable to the situation in order to meet the growing expectations of inclusivity from its diverse end-users. This will be dealt with in the personalisation work to be undertaken after the delivery of the basic avatar design.

2.3. State of the Art of Virtual Characters Design

As discussed in Llorach et al.[6], the most important aspects of character design to achieve conversational characters are: high-level geometry control; detailed geometry control; clothes, hairstyles; textures; facial blend shapes and skeleton; and realism. Tools to create or adapt characters can save much time, and a comparative table of tools including these aspects is presented in the paper (p. 124). Our design of the avatar has taken these aspects into account, adding careful eye design also. There are other important requirements, also discussed in Llorach et al., which are non-functional requirements derived from characters having to be web-based, related to file size to meet acceptable downloading times and real-time rendering; facial expressions based on a reduced number of blend shapes or facial bone rigs for quick rendering; and similar aspects for skeleton and animations.

A significant challenge in terms of users' expectations comes from the recent initiatives related to entertainment, the most significant initiative being the MetaHuman Creator[7] launched by Epic Games, which provides virtual humans of unprecedented quality to be used in real-time games (higher quality virtual characters mostly for cinema is the target of the PRESENT EU project with the participation of UPF-GTI, which also involves Epic Games UK). DHH expectations will probably be very sensitive towards signing avatars below this level of quality. During the course of the project, SignON will aim at reaching this quality level, and in the second half of the project, the use of MetaHumans will be explored.

[6] Llorach, G., Agenjo, J., Blat, J., Sayago, S. (2019). Web-Based Embodied Conversational Agents and Older People. In: Sayago, S. (ed) *Perspectives on Human-Computer Interaction Research with Older People*. Human–Computer Interaction Series. Springer, Cham. https://doi.org/10.1007/978-3-030-06076-3_8

[7] https://www.unrealengine.com/en-US/digital-humans

As an indicator of the level of quality of MetaHumans, we reproduce an image taken from the MetaHuman Creator web in Figure 3.

Fig 3: Example of MetaHumans avatar quality

2.4. Manual/Non Manual Features of Sign Languages and Design

As discussed earlier in Section 2, the basic character to be designed needs to be able to support sign animations rather than the more typical locomotion animations related to entertainment. Hand animations within the MFs require a more fine grained skeleton than usual, and the design should take this into account. Similarly, the blendshapes/facial bone rigs should also meet the demands for high quality NMFs, supporting expressivity and mouthing. Figure 4 illustrates some of the challenges of hand animation in the context of SLs, namely those of the precise and sophisticated positions required.

Fig 4: **Left**: Representation of expressivity in SL[8]. **Right**: Signs of the American Sign Language alphabet[9].

[8] https://www.quartoknows.com/blog/quartoexplores/learn-a-new-language-asl-for-beginners
[9] https://analyticsindiamag.com/hands-on-guide-to-sign-language-classification-using-cnn/

3. Skeleton/Rigging

The animation of *articulated structures* such as persons is usually based on *skeletons*, i.e. hierarchical structures of segments (bones) and joints. Specifically, rigging refers to the process of creating the bone structure (skeleton or armature) of a 3D model. This bone structure is used to manipulate the 3D model like a puppet for animation. The definition of the skeleton of the virtual character has been considered mainly in two sources related to SLs: the specifications from Irene Murtagh's doctoral thesis[10], and the work on the creation of a corpus for the French SL (LSF) based on Motion Capture[11]. On the other hand, the capture of our end-to-end system (discussed in the companion deliverable D5.8 *A Realiser of BML-based Script to 3D Animated Character*) is based on MediaPipe, which uses a different representation of a skeleton from the one we mostly use, and we need to take that into account.

3.1. Sign Language Supporting Skeleton

In Murtagh's doctoral thesis, an armature is proposed that is able to animate signs. The 62 skeleton joints are distributed as follows:

- 2 joints in the head.
- 4 joints in the spinal column.
- 4 joints in both legs.
- 4 joints in both feet.
- 6 joints in both arms.
- 42 joints in both hands.

Figure 5 and 6 show the bone distributions and the avatar joint hierarchy, respectively, for the MF.

[10] Irene Murtagh: A Linguistically Motivated Computational Framework for Irish Sign Language, PhD thesis, Centre for Language and Communication Studies, University of Dublin, Trinity College, 2019.
[11] Naert, L., Larboulette, C. and Gibet, S.,: LSF-ANIMAL: A Motion Capture Corpus in French Sign Language Designed for the Animation of Signing Avatars. In *Proceedings of The 12th Language Resources and Evaluation Conference* (pp. 6008-6017), 2020.

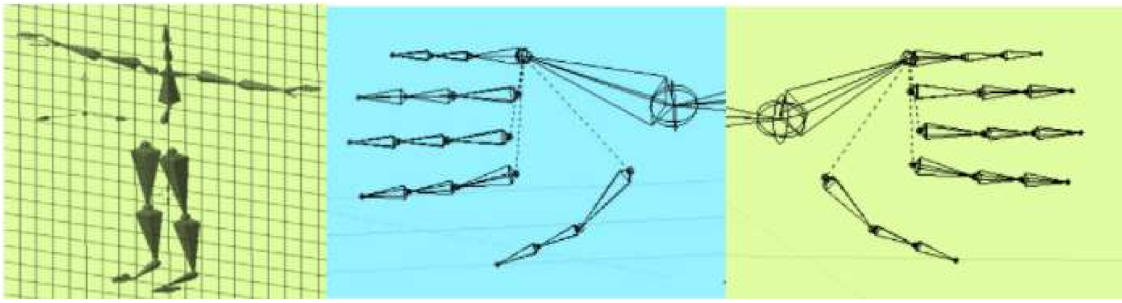

Fig 5: Armature defined in Murtagh's thesis[12]

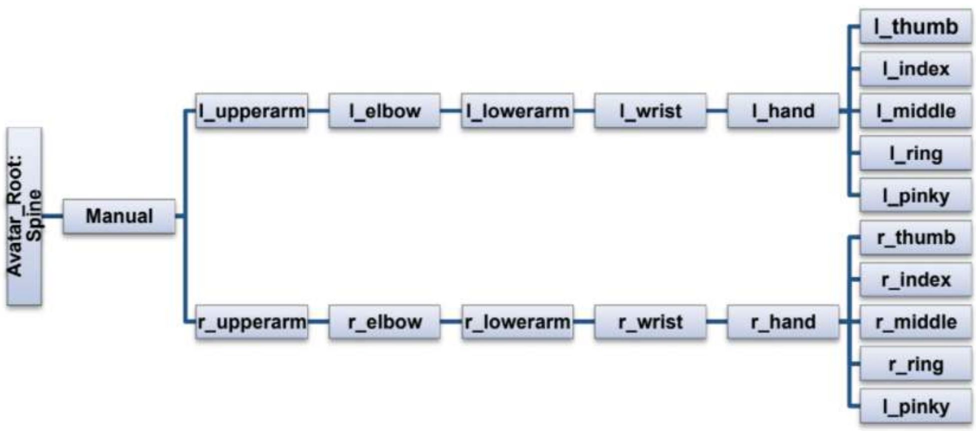

Fig 6: Joint distribution used in the manual features

The joint hierarchy is also provided for the NMF (face and shoulders). However, our definition of the mark-up language works without a direct relation of our input and the bones, therefore we have more freedom to design the face bones as desired.

If we compare it with our current model (Eva), we can appreciate the similarities. The armature of Eva (shown in Figure 7) has 67 joints distributed as follows:

- 4 joints in the head.
- 5 joints in the spinal column.
- 4 joints in both legs.
- 6 joints in both feet.
- 6 joints in both arms.
- 42 joints in both hands.

[12] Irene Murtagh: A Linguistically Motivated Computational Framework for Irish Sign Language, PhD thesis, Centre for Language and Communication Studies, University of Dublin, Trinity College, 2019.

Fig 7: Armature of the current model (Eva). Hands in detail (left), whole body (right).

3.1.1. MediaPipe Data

We also need to consider the technology that we use to capture data and the joints it outputs. The output is not actually in terms of joints because it does not involve a virtual character, but instead the locations of *pose landmarks*. In fact, MediaPipe pose tracker gives the location of 33 pose landmarks (see Figure 8):

- 11 in the head.
- 0 in the spinal column.
- 4 in both legs.
- 6 in both feet.
- 4 in both arms.
- 8 in both hands.

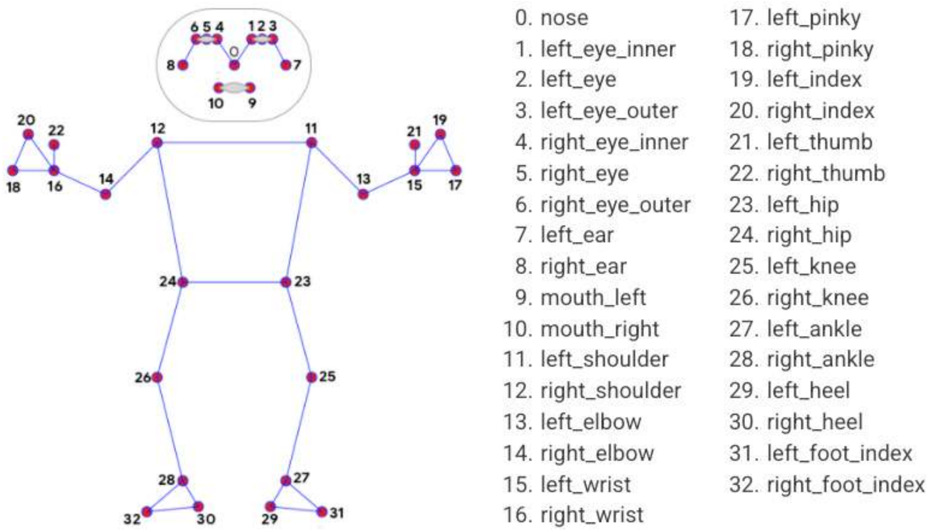

Fig 8: Landmarks distribution for the pose[13]

Moreover, MediaPipe's hand tracker gives the keypoint location of 21 3D landmarks, as illustrated in Figure 9).

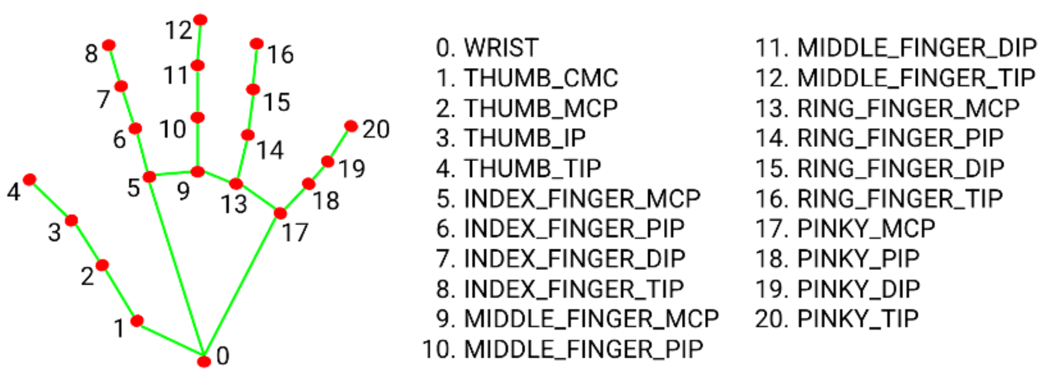

Fig 9: Landmarks distribution for the hands[14]

3.1.2. Reference Pose

For the reference pose, we use a T-Pose which complies with the specifications below.

- The arms must be spread along the X-axis. The left arm should therefore be pointing along the positive direction of the X-axis.
- The top of the character’s head must be up, in the positive direction of the Y-axis.
- The character’s hands are flat, palms facing the ground, with the thumbs parallel to the X axis.

[13] https://google.github.io/mediapipe/solutions/pose
[14] https://google.github.io/mediapipe/solutions/hands

- The character’s feet need to be perpendicular to the legs (with the toes pointing along the Z-axis). The feet must not be rotated around the Y-axis (meaning the toes of the left foot should not point inward toward the right leg or outward away from the right leg).

The 3D model and the armature need to be in the same position during the rigging process. By doing this, we can relate the movement of each bone of the skeleton with the vertices of the 3D model. Our end-to-end system, discussed in the companion deliverable D5.8, allows for the exchange of skeletons, and thus, we do not discuss this further here.

4. The Character Beyond its Skeleton

4.1. Eva

Eva already was accepted by most of the community working with SignON, thus our basic character for the project stems from further work on it in order for the character to fit the purposes of this project. We next describe Eva's characteristics.

Eva is a model used in previous projects (KRISTINA, H2020 G.A. 645012, among others) and ongoing ones (PRESENT, H2020 G.A. 856879) as an embodied conversational agent. She was created from a free scanned model available on Adobe Fuse. She has a skeleton with 65 bones, 52 morph targets (different deformations of the same mesh) and different textures with good resolution. She was considered a good model because she looks realistic, but not *too* realistic, thus possibly avoiding the Uncanny Valley. From that base, improvements were made both in the rendering and for the animation. Some key technical and aesthetic traits were as follows: The character was intended to look empathic and yet authoritative. The future work on customisation should address other aspects such as ethnicity and gender. The facial rig supported enough expressivity: bones provide additional mobility while the blendshapes are necessary so that the avatar can take the necessary states to be expressive and support lipsyncing. Figures 10 and 11 show some of Eva's poses:

Fig 10: Visualisation of our current 3D character (Eva)

Fig 11: Different views of (old) Eva

In the next subsection we discuss the improvements of the character for SignON.

4.2. A New Character

With the hardware improvements since Eva was created, a more realistic model can be supported even in low-end devices. Therefore, a new virtual character model has been created.

Some significant improvements are on the face proportions and the detail of the hands, and by introducing a proper physically-based rendering (PBR) of the model by using production-level maps/textures, the necessary PBR information is provided. The asset generation counts with the support of an artist, who follows the requirements which have been provided, reflecting the previous content of this deliverable.

Efforts have been invested into a better rendering of hair and eyes, which convey a lot of visual quality and are considered very important to users' satisfaction. Both hair and eyes usually need to be rendered through specialised shaders (*shaders* are GPU-based programs, used in computer graphics).

The skeleton and rigging needed to be updated, so that the character could support SLs, both in terms of capture, rendering, and potential interoperability. The hands are very important for SLs, and we have been inspired by previous work by other teams working on SL avatars. We have also been dealing with some complexities caused by using deep learning based software for capture. Both latter aspects were discussed in Section 2.

Figure 12 is an initial design of the new Eva. As we can see, it has more details in the face, and the hair was modified as we adjusted our requirements to maximise the quality of the model and the animations.

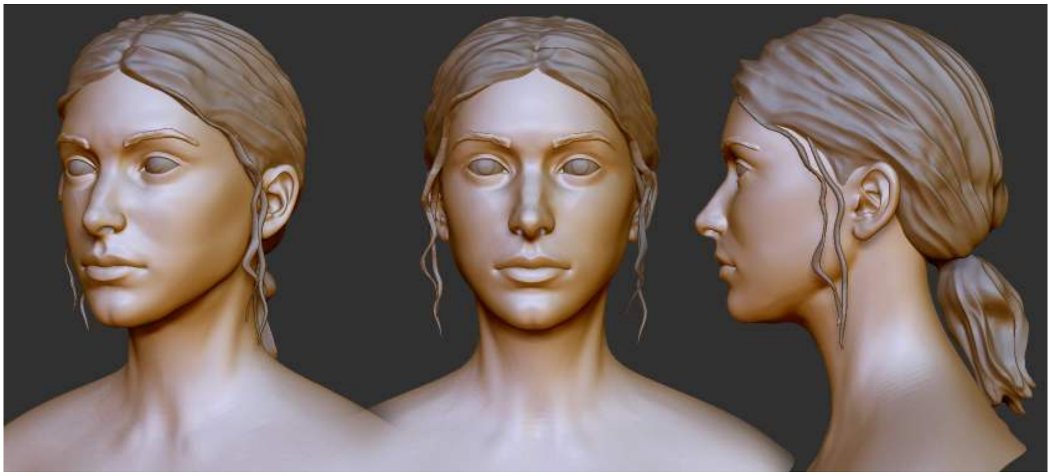

Fig 12: Initial design of new Eva

Figures 13 and 14 are the final meshes for the face, eyes, and hands:

Fig 13: View of the face mesh (no hair, no eyes) (left), and of the eye mesh (right) of the new Eva

Fig 14: View of the new hand mesh of Eva

4.3. Blendshapes

Blendshapes (or Morph Targets) are key elements on conveying facial emotions and non-manual features in general. Not only are they simple to implement and understand (even for non programmers) but they can easily parameterise facial expressions into a small set of values.

Each blendshape consists of a mesh that contains the same vertices but in different positions and normals. A base/neutral pose is also required. The difference between a weighted blendshape and the neutral pose can be computed and then applied to the neutral pose. By doing this procedure over all blendshapes, transitions among poses can be rendered without visual glitches. Once the final pose is computed, the rendering can be applied as it were for a single mesh. This is illustrated in Figure 15.

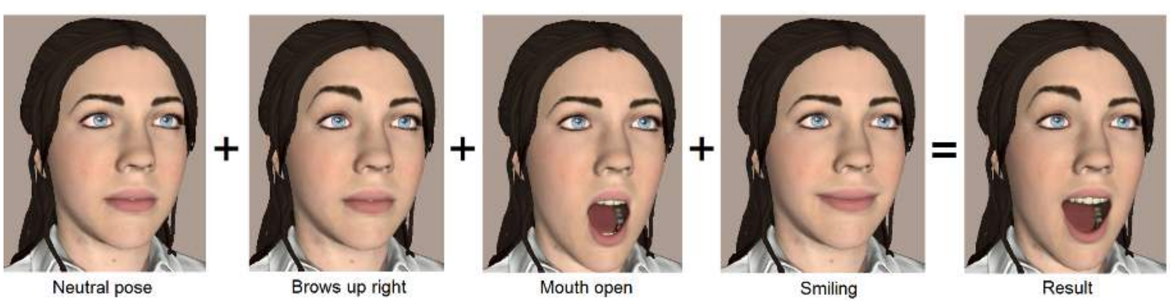

Fig 15: Example of blendshape addition. Retrieved from D5.8, figure 15

Each blendshape may be targeted at a specific region of the face, keeping all other vertices in neutral pose. This way, the different procedures and expressions the face is involved in can be computed semi-independently without much complexity. For example, there is a set of blendshapes in the Eva avatar dedicated exclusively to the mouth (lips, jaw aperture and tongue). These are used by the mouthing module. Another module is in charge of blinking and other procedures are related to eyelid movement. Both modules can work separately and apply changes directly to the weights of the blendshapes they need.

The Eva model used up until now lacks some blendshapes and some of the existing ones need urgent improvement. For example, the tongue is a muscle that is constantly moving and changing its shape. In certain mouth utterances (either from mouthing or SL gestures), the widening or the degree to which the tongue is out of the mouth is relevant. However, the current model only contains a basic "front of the tongue upwards" morph target that feels very clunky. Other blendshapes related to the mouth also require some improvement such as the narrowing of the lips (i.e. a kissing shape) or the lips being closed tightly (as when pronouncing a /p/).

Since a new character will be created, all aforementioned morph targets fixes and additions can be directly included as the modelling requirements, which will result in more natural poses rather than fixing the old model. The complete list of all blendshapes that are required can be found in Table 1 in Annexes.

4.4. Eyes

Eyes can significantly influence the expression of the avatar. In that sense, it is important to replicate their visualisation as much as possible. In the proposed material, cornea refraction and wetness are taken into account, plus proper customisation options, like iris colour or iris radius.

For a proper blending of the eye iris and sclera, two separate textures are required, one with information from the iris and another one with information of the sclera. This is important because both components will usually have different properties, like would happen in a multi-layered material.
Such layers are exposed in Figure 16. Using a grayscale iris allows for faster customisation of the iris colour when defining the material.

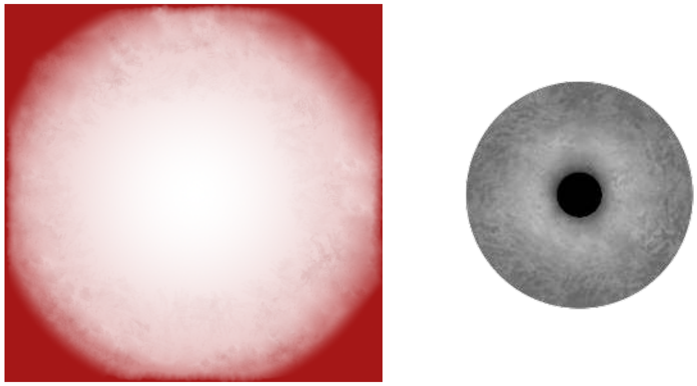

Fig 16: Examples of iris (right) and sclera (left) in separated textures

To account for a physically-correct cornea refraction, the eye model should have a frontal bulb, as shown in Figure 17.

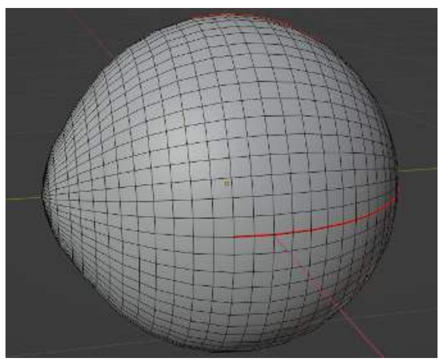

Fig 17: Example of a 3D model of the eye

Also, the model size should be normalised, or at the very least, the vertex position projected into the frontal plane must match the textures UV (in the range [0,1] or [-0.5,0.5], as exposed in Figures 18 and 19.

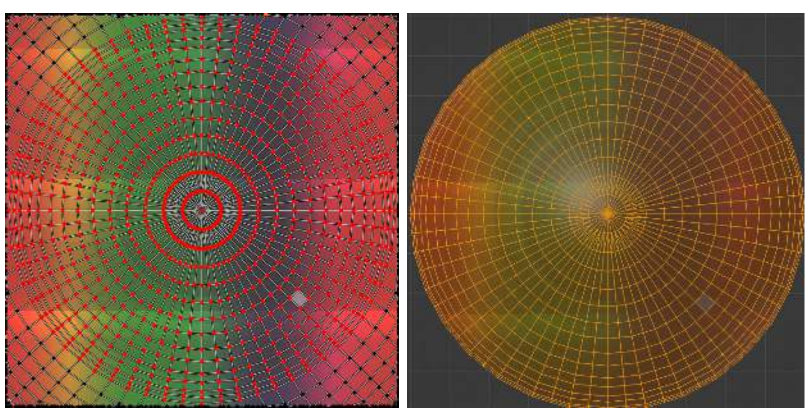

Fig 18: UV Layout

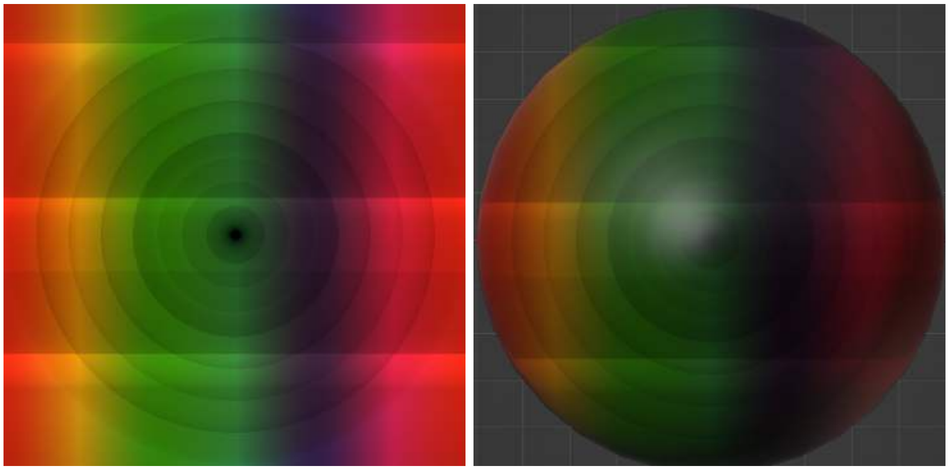

Fig 19: UV Layout

With the proper layout, the frontal orthographic view of the model should preserve directions and size with respect to the texture data. The resulting eye visualisation, with refraction and a multi-layer setup taken into account, achieves a realistic simulation of the lighting phenomena happening in the eye (see Figure 20).

Fig 20: Example of a final render of an eye

4.5. Hair

Hair is a complex component in real-time rendering. The solution for this case, as done in most interactive applications, is to use a hair cards model. Hair cards are a set of planes with baked strand information on textures. If blended properly, a good representation of the hair volume can be achieved.

As for texturing, a different set of maps are required, such as tangency, strand ID, and opacity maps. Tangency is used for hair-specific shading computations. For simpler shading complexity, this could be substituted with anisotropy. Strand ID is used to assign small colour variations to each strand, achieving a better per-strand detail and thus removing uniformity across a single hair card. Opacity will be used to alpha test each fragment, thus achieving a more realistic representation of hair at different depth layers.

The desired result would be similar to Figure 21.

Fig 21: **Top:** resulting hair visualisation. **Bottom:** Hair cards geometry (before texturing/shading). (Retrieved from [15])

With respect to the state of the new Eva, there is a small delay in the creation of the hair. We have planned to have it finished by July.

4.6. Texturing

As stated in Section 4.2, multiple textures must be provided per material in order to account for physically-based surface reflection. Each texture is created specifically to give information about their corresponding 3D object details in a compressed way. Therefore, using more textures will increase the complexity of the render but will also improve the quality of the final result. For our previous Eva we

[15] https://80.lv/articles/creating-hair-for-real-time-game-characters/

used 4 textures (albedo, normal, roughness, subsurface scattering), while in the new model we currently have 8 textures. We list the existing textures for the skin (face and hands) of the new Eva with a brief explanation for each:

- The **Albedo** contains the base colour of the 3D model. It defines the diffuse colour of the surface.
- The **Normal** (bump map) gives the depth information of the 3D model. It is used to compute how the light is reflected off the 3D model.
- The **Roughness** (gloss map) defines how light scatters across the surface of the 3D model. Its value can generate a range of surface finishes, from shiny, like plastic, to matte.
- The **Ambient Occlusion** is used to provide information about realistic global shadows. This texture saves a large amount of computations on the render time.
- The **Subsurface Scattering** is used to simulate the effect of light scattering inside a 3D object. As can be seen in the texture, it is very pronounced in important places for our avatar (eyes, hands, ears, mouth).
- The **Cavity** stores small-scale ambient occlusion. The addition is that it captures smaller details.
- The **Curvature** stores the convexity/concavity of the mesh. It can be used to give details about the wear of the surface or to indicate where there might be dirt accumulation.
- The **Thickness** map measures how thick/thin the surface is. It is used in lighting computation.

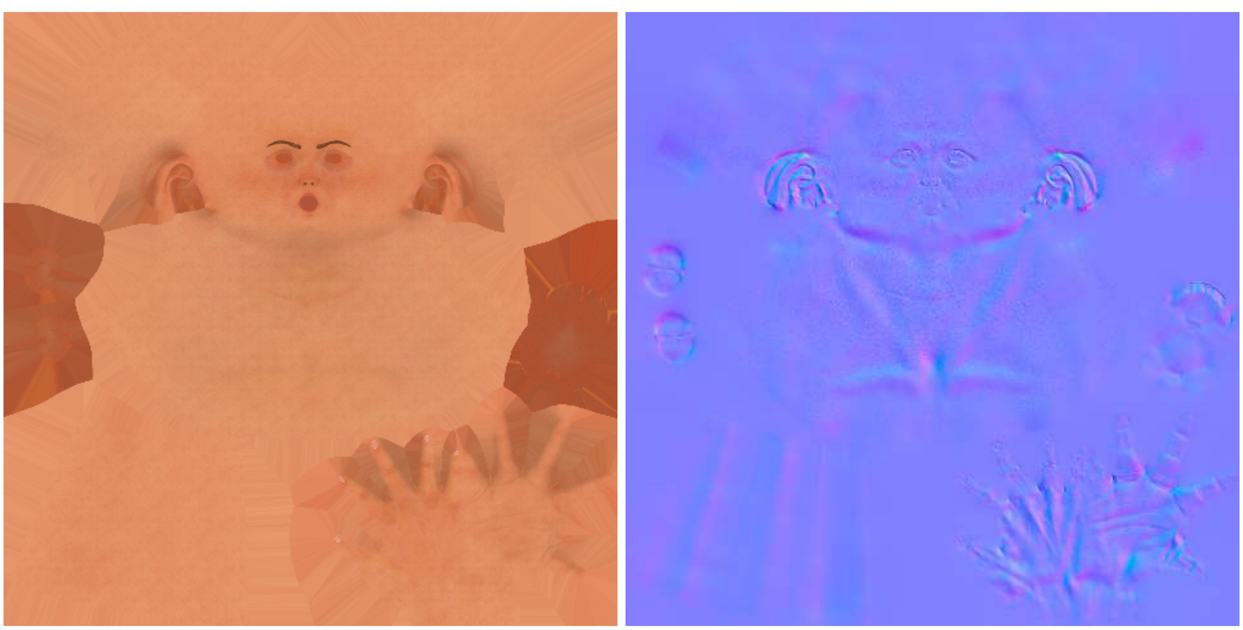

Fig 22: **Left:** Albedo texture. **Right:** Normal texture.

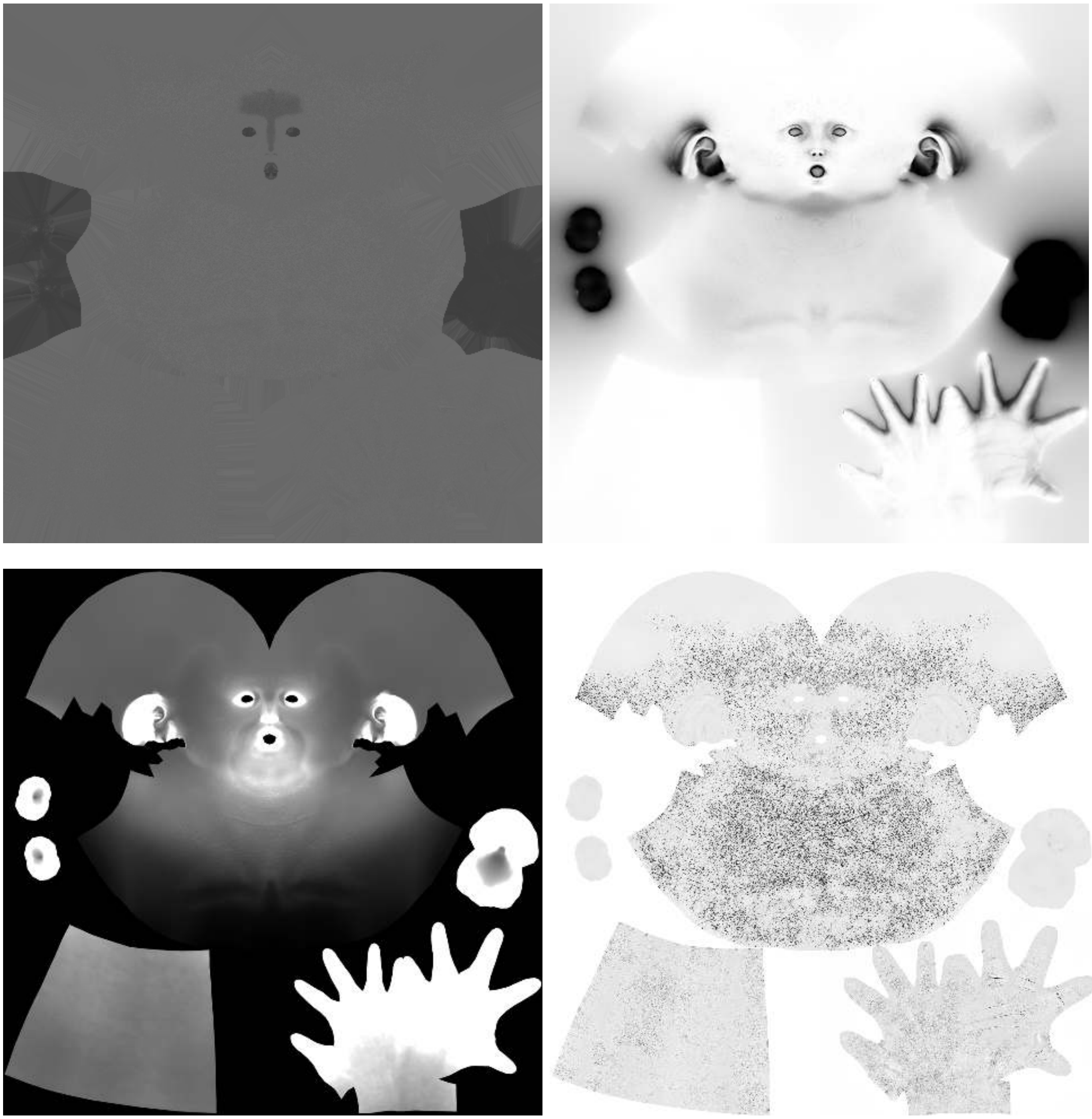

Fig 23: **Top-Left:** Roughness texture. **Top-Right:** Ambient Occlusion texture.

Bottom-Left: SubSurface Scattering texture. **Bottom-Right:** Cavity texture.

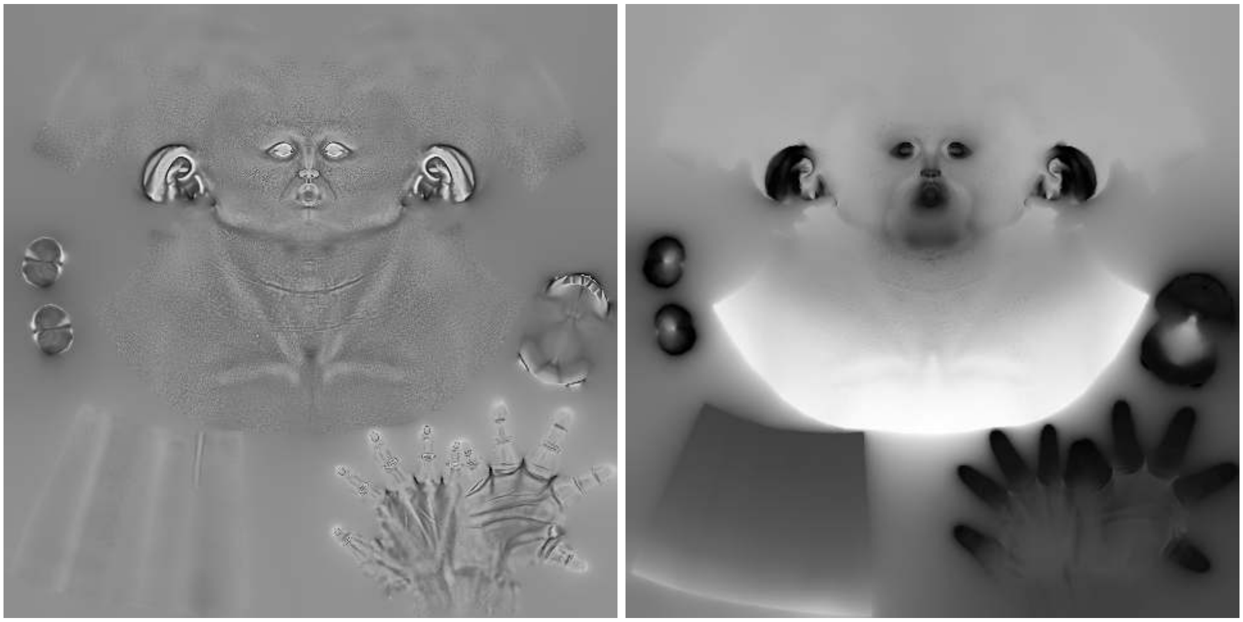

Fig 24: **Left:** Curvature texture. **Right:** Thickness texture.

As it can be appreciated in Figures 22, 23, and 24, these textures are only for the skin. Eyes and hair require particular input textures. Additionally, the textures for each 3D model differ. For the eyes of the new Eva we created 4 textures (albedo, normal, roughness, and specular) for the iris, and the same 4 textures for the sclera. The requirement to separate the iris from the sclera is explained in the following section. We can see the textures in Figures 25 and 26:

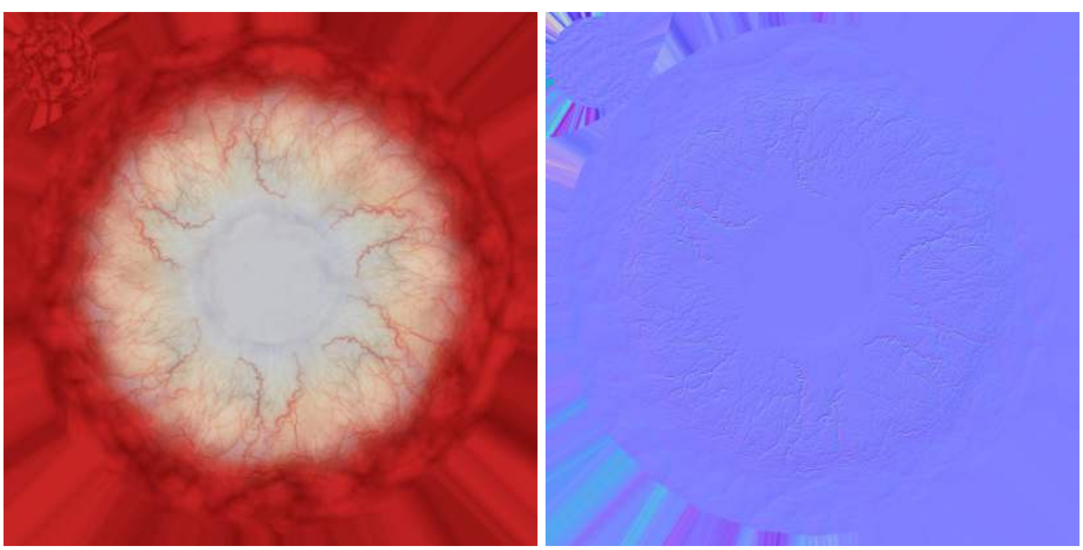

Fig 25: Esclera textures: **Left:** Albedo texture. **Right:** Normal texture.

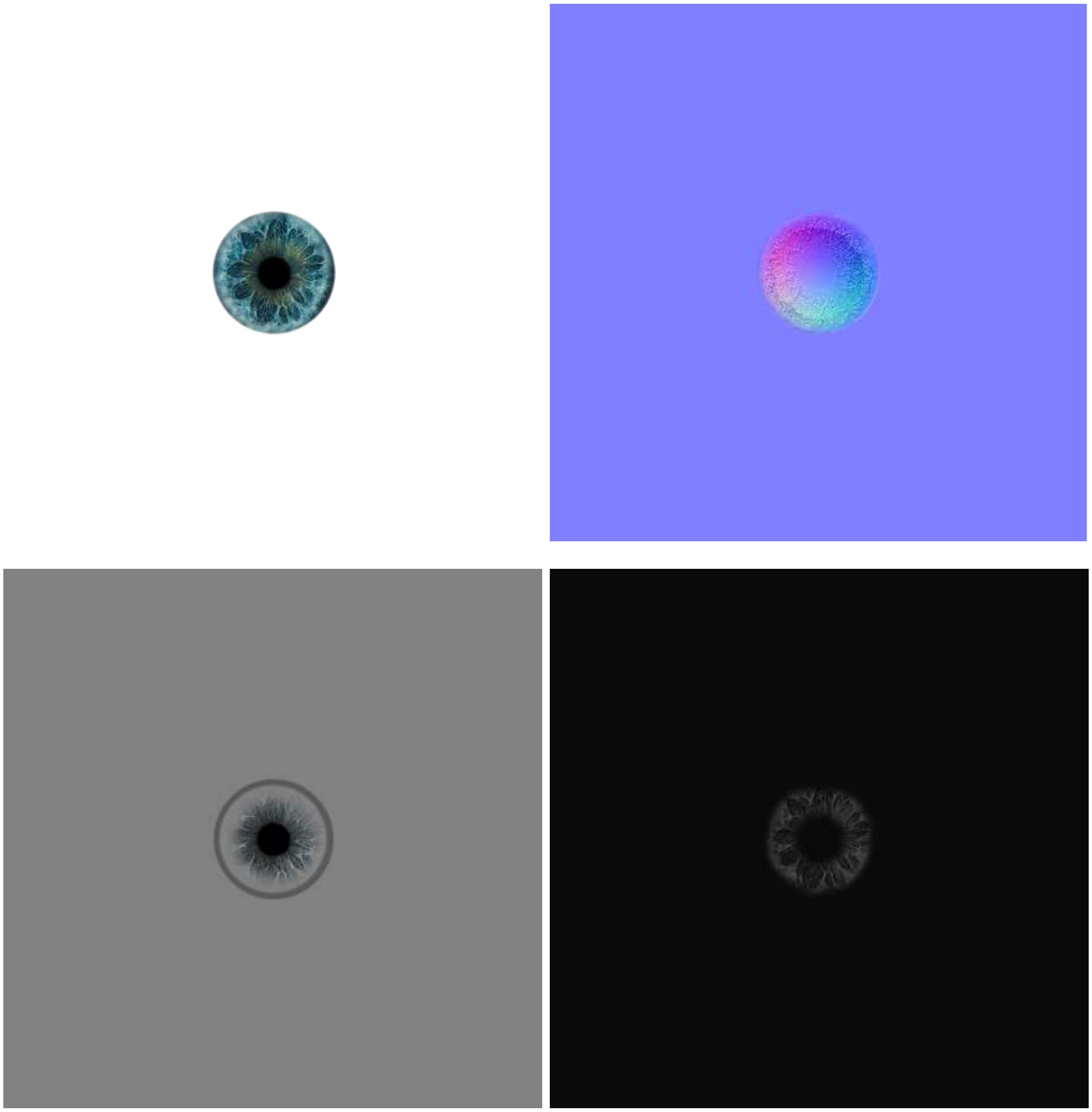

Fig 26: Iris textures: **Top-Left:** Albedo texture. **Top-Right:** Normal texture.
Bottom-Left: Roughness texture. **Bottom-Right:** Specular texture.

For the hair, we created 5 textures (albedo, specular, normal, ambient occlusion, and transparency). Note that the hair will also contain the eyelashes, so it makes sense that they need a transparency (or alpha) texture, which is not needed in the skin nor the eyes.

5. Conclusions

In this document we discussed the technical aspects of the new model of virtual character created for the SignON project, providing the background which has motivated this design. Its essential aspects are a more detailed skeleton which can be used for SL animation, and significant improvements in the visual quality, which should help increase the acceptability of the virtual signer, in the context of the quality of avatars commercially available. The new model is ready, with some slight delay in the hair textures, which will be available in July 2022.

The next deliverable in month 36 will include specific explanations for the personalisation of avatars. Including diversity aspects like gender, skin colour; and technical aspects like the clothes worn by the avatar and the different hairstyles.

6. Annexes

Body Part	Lexeme	Action Unit[16]	Blendshapes name used
Eyebrows	FROWN	AU4	brows_down (L+R) + brows_in (L+R)
Eyebrows	ARCH	AU1	brows_outer (L+R) + brows_up (L+R)
Eyes	OPEN_WIDE_EYE	AU5	eyes_wide (L+R)
Eyes	SQUINT	AU44	squint (L+R)
Eyes	BLINK	AU45	blink (L+R)
Eyes	CLOSED	AU43	blink (L+R)
Cheeks	SUCK_IN_RIGHT	AU35 right	new (1)
Cheeks	SUCK_IN_LEFT	AU35 left	new (2)
Cheeks	SUCK_IN_BOTH	AU35	new (1+2) (L+R)
Cheeks	BLOW_RIGHT	AU33 right	cheek_puff R + new (3)
Cheeks	BLOW_LEFT	AU33 left	cheek_puff L + new (4)
Cheeks	BLOW_BOTH	AU33	cheek_puff (L+R) + new (3+4) (L+R)
Mouth	OPEN_WIDE_MOUTH	AU27 + AU14	mouth_open + new (5)
Mouth	CLOSE_TIGHT	AU24	new (3+4)
Mouth	SMILE_TEETH	AU12 + AU25	smile (L+R) + lower_lip_out + mouth_open
Mouth	SMILE_TEETH_WIDE	AU12 + AU25	smile (L+R) + lower_lip_out + mouth_open
Mouth	SMILE_CLOSED	AU12	smile (L+R)
Mouth	ROUND_OPEN	-	new (6) (com la O)
Mouth	ROUND_CLOSED	AU23	mouth_narrow (L+R)
Tongue	OUT_POINTED	AU19	tongue_out
Tongue	OUT_ROUND	-	new (7) (?)

[16] https://imotions.com/blog/facial-action-coding-system/

Nose	CRINKLE	AU9	nose_scrunch (L+R)
Nose	FLARE	AU11	new (8)
Tongue	-	-	tongue_wide (new 9)
Tongue	-	-	tongue_back_up
Mouth	-	-	frown

Table 1: Lexeme - Blendshape relation for Eva