Fast creation of look-a-like avatars

Xenxo Alvarez[1], Anna I. Bellido Rivas[2], Nuno Barbosa[1], Kristopher J. Blom[2], Bruno Oliveira[1], Verónica Orvalho[1]

[1]Porto Interactive Center, Departamento de Ciência de Computadores, Faculdade de Ciências, Universidade do Porto, Rua do Campo Algre 1021, 4169-007 Porto

[2]Universitat de Barcelona, Facultat de Psicologia, Departament de Personalitat, Avaluació i Tractaments Psicològics, Passeig de la Vall d'Hebron 171, 08035 Barcelona, Spain

xenxo3d@gmail.com, nuno.barbosa@fc.up.pt, blom@ub.edu, bruno.oliveira@dcc.fc.up.pt, veronica.orvalho@gmail.com, abelliri8@alumnes.ub.edu

Resumo

Uma sensação de imersão é o principal objectivo dos ambientes virtuais, como os videojogos.

O sentimento de "estar" no ambiente pode ser melhorado através do uso de avatares que sejam semelhantes ao participante, ou participantes. Essa semelhança virtual é possível de criar nos días de hoje, mas trata-se de um processo lento e para qual é necessário um especialista nas áreas de modelação e controlo de personagens. Estes requerimentos tornam a introdução rápida de novos participantes num ambiente virtual practicamente impossível.

Apresentamos um processo através do qual um leigo é capaz de produzir um avatar virtual que apresente semelhança a um qualquer participante, em apenas 30 minutos. Faz uso de software comercial e de hardware comum, o que permite que este processo possa ser replicado em qualquer lugar.

O modelo criado tem um controlo de animação completo, podendo ser imediatamente animado, quer através de captura de movimento, quer por um processo manual. O controlo de animação foi pensado para reproduzir animações subtis, importantes para o sentido de imersão.

Neste artigo apresentamos o processo e os desavios encontrados para o criar.

Palavras- chave: Ambientes Virtuais, Aparência Virtual, Processos de Criação, Computação Gráfica

Abstract

Providing a better immersion sensation is one of the main objectives of virtual environments, like videogames.

The feeling of being in the environment can be improved by having avatars that look like the participant, or participants. Such virtual likeness can nowadays be achieved, but this takes a considerable amount of time to create and an expert in the field of modeling and rigging. These conditions make it impossible to quickly introduce new participants in a virtual environment.

We present a pipeline with which a non-expert can create a virtual avatar that resembles any participant, in just 30 minutes. It uses commodity hardware and software, allowing the pipeline to be replicated anywhere.

The produced model is also fully rigged, which permits imediate animation, either by hand or by motion-captured data. The rig was developed with subtle animations as a goal to grant the avatar the possibility to replicate small movements, important for the feeling of immersion.

Here we describe the pipeline and the challenges we found to create it.

Keywords: Virtual Environments, Virtual Likeness, Pipeline, Computer Graphics

Introduction

Virtual environments, like videogames, offer the opportunity to immerse one or more individuals in a completely different virtual world with which they can interact, and an all-new of experience is open to participants today, when compared to the inception of virtual reality. The photorealistic rendering techniques that have been developed along the years, the possibility to do realistic physics simulations and the democratization of high-end virtual reality gear, like head-mounted displays, allow nowadays a complete immersion sensation, in which anyone can fell integrated with the virtual environment, undergoing the same level of interaction and sensations as if in the real world.

One of the main drawbacks of this immersion sensation is visual likeness. If a participant sees himself in the mirror, for instance, he must believe that the virtual character he is looking at looks and acts like him, otherwise the immersion sensation will be lost.

Although possible to achieve with today's technology, the process to produce a look-a-like character is tedious, laborious and requires an expert in the field of virtual characters creation. It starts with data acquisition to get the shape of the face and the texture. The data comes, most frequently, from a facial 3D scanner and has to be cleaned from noise and incorrect topology. Once cleaned, the data is used to model a facial mesh. After the facial mesh is correctly created, a process called rigging follows, in which the mesh is set for animation, with the inclusion of a control structured. This rig has to be prepared for real-time motion captured animation, since the virtual character has to mimic all movements from the participant instantaneously.

At the end of this process, a fully animatable model of a human character will be available, but it took, approximately, 4 weeks of work of an expert artist. To be able to provide to any participant an enriching experience in a virtual environment, there is the need to develop tools that can create in a short period of time a look-a-like character of a particular human, which can be used by anyone, not necessarily an expert in the field.

With this in mind, we developed a proof-of-concept that can produce a completely animatable character, which looks like a given human, in less than 30 minutes. This was developed within the context of a research project focused more on facial and body likeness than on a precise representation regarding cloths or hair. Taking advantages of

well-known software and facial and body meshes rigged by default, we devised a pipeline that takes as input a couple of photos — front and side — and produces a animatable avatar — both face and body — that can be used in a virtual environment and mimics the likeness of the participant. We also developed a set of tools that allows any user to create the avatar without the need of understanding the details of the avatar's creation process.

Pipeline

The pipeline (Figure 1) to produce an animatable avatar is composed of 3 stages: facial pipeline, body pipeline and integration.

In the next sections we will depict each one of the stages of the pipeline. We will start by describing the common steps shared by the facial and the body pipelines, followed by a thorough description of each stage, in which we will explain the process and the software involved.

Development stages

Both the facial pipeline and the body pipeline require three different steps, the same but adapted to each scenario: modeling/texturing, rigging and motion capture data integration.

The modeling/texturing step refers to the creation of the meshes that correspond to the face and the body of the avatar. The rigging step corresponds to the creation of controls that allow animating the avatar. Motion capture data integration corresponds to the mapping to the avatar of the collected motion capture data. This will allow to animate in real-time the virtual character based on the input given by the participant's motion — both facial and body motion.

Facial pipeline

The modeling/texturing stage in the facial pipeline takes advantage of the FaceGen (Singular Inversions, 2009) (Figure 4) software to generate a model of the face that resembles the participant's facial features and shape. It is based on two to three input photographs of the participant's face – one front view and two left views of the face --, which are used to create a fitting between the photographs and a default facial mesh, and also to extract the texture of the face.
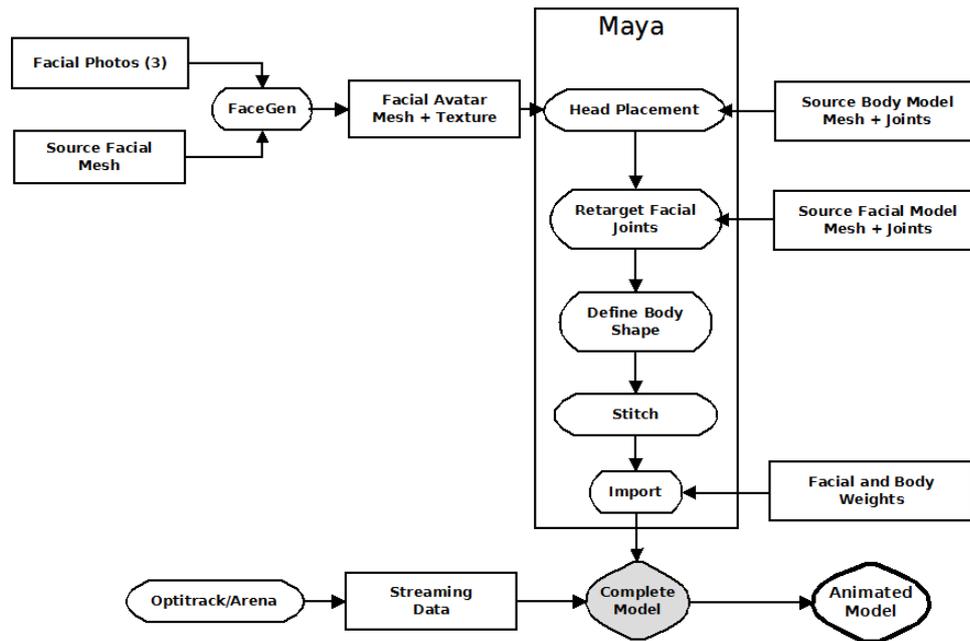
Figure 1: The complete pipeline

The original mesh provided by FaceGen presents some handicaps that we had to tackle, in order to render the face in real-time and to have a seamless integration with the body. Most of these problems relate to the mesh's topology. With over 12000 triangles, although a good polygon count for fitting, it makes rendering in real-time a difficult task. Another problem was an overlap of vertices in the mouth region, which can cause odd behavior in lip animation. Finally, the UV map for textures also had overlaps, which adds an extra effort in the case we want to switch the face's texture.

Furthermore, the mesh generated by FaceGen is not rigged, and a rig for it would have to be created. This rig has to give a comfortable level of control over the facial animation, and, for this, we would have to be able to access the core mesh of FaceGen, to make the rig as generic as possible in order to simplify the transfer later in the pipeline.

To address all these issues, and another one related to stitching, which will be address later on, we opt for doing the process in reverse. FaceGen provides a tool for using custom made facial meshes for the facial fitting algorithm. This tool, called FaceGen Customizer (Singular Inversions, 2009), is an add-on to the program and is available separately from the main software. For this, we created a generic rigged mesh, with a correct topology and an acceptable triangle count, and provided it to the Customizer. Customizer takes around 6

hours to complete, but this process only has to be run once. As soon as the mesh is ready for FaceGen, no further processing is required.

Besides fitting the available topology to the provided photographs, now to a custom facial mesh, FaceGen also samples among these to create a facial texture that mimics the user's facial color. Because it is a sampling process and not a direct match, facial features like moles are discarded, but the overall result is good enough to create a virtual likeness sensation. With FaceGen, the facial generation process is almost automatic, and the only manual intervention is on landmarking, to match features such as eyes, facial extremes, etc, but it can be done by any person, not necessarily an expert.

Once the facial mesh is generated, the second step is to take the facial model and assign to it a rig. This rig will allow manipulating the facial mesh in order to produce facial expressions. These facial expressions will be drive by the motion captured data from the user and thus have to be as realistic as possible. To achieve this, the rig must be carefully placed and present a great level of control over the mesh. The chosen rig to animate the generated avatars is composed of 37 joints, and the controls where placed according the deformation areas of the face. Areas with more deformation, like mouth, eyes and cheeks have more density of control points than those who do not present much deformation, like the nose.

The automatic rig of the facial mesh is done using the TPS algorithm (Orvalho, Zacur, & Susin, 2008). This algorithm automatically transfers a rig from one mesh to another, maintaining the rig's characteristics — placement and influence of each joint. As previously mentioned, a generic rigged mesh was created, and it is this rig that is then transferred to the new generated mesh.

To transfer the rig in the facial mesh, the commercial modeling tool Maya (Autodesk, Inc, 2010) is used. Maya was originally used to create the generic rigged mesh, and the transfer algorithm also uses Maya's support to ease the process.

The transfer is automatic, with minimal user intervention. Our team created a script using the scripting features in Maya — by the means of the Python language — that automatically applies the generic rig to the generated facial mesh (Figure 4).

Once the face model is prepared, having a look that resembles the participant's face, and rigged, it is possible to map, in real-time, motion captured data to this same model, so that the virtual character mimics the movements made by the participant's face.

In our tests, we used the NaturalPoint's Optitrack (NaturalPoint, Inc, 2010) motion capture hardware. NaturalPoint's solution for facial motion capture is a set of, at least, 6 infra-red cameras that capture reflective markers placed in the participant's face. The number of markers is arbitrary, but a commitment between the accuracy of the capture and the number of markers must be established. Our test shown that the minimum number of markers to achieve the level of detail required to induce a likeness feeling to a user is 37. After retargeting, to match the facial measures of the participant's face to the virtual face, a one-to-one mapping is done between the markers' position and the joints in the previously mentioned rig.

Body pipeline

As with the facial pipeline, the body pipeline starts with a generic rigged body mesh that can be adjusted to the anthropomorphic measures of the participant. In order to do this adjustment, a number of body shapes for the same model where created in Maya, and then a simple morph-based method is used. Whether the user is tall, short, thin or strong, the number of shapes created allows approximating the model's shape to the participant's shape.

The base model was also rigged with 57 control points, which allows a seamless integration with the body motion capture system, mimicking the participant's movements in real-time. Each joint in the model is also scaled to match participant's physiognomy, and this process is done automatically by the motion capture data.

Integration

Integration refers to process of stitching the head to the body, creating, this way, a fully body animatable avatar. The analogy with stitching is almost perfect, since the facial mesh is connected, polygon-by-polygon, to the body, in a process that resembles most stitching an open wound.

This stitching process was another reason to use a custom mesh with FaceGen. This allows us to control the number of polygons in the neck area, both from the facial mesh and from the body mesh, which renders the process almost automatic. If one has the

control over the vertices, and placement of these, even if some scaling is necessary to match the size of the head to the size of the body, the stitching becomes trivial, and can be completely automated.

To achieve this level of control, not only a facial mesh was modeled and rigged, but a full body was created. The head was then chopped of, and used in FaceGen Customizer. Since neither the facial pipeline, nor the body pipeline adds or removes vertices, a perfect match is granted in the stitching process.

Again, the tools created in Maya ensure an automatic stitching process, creating a complete rigged model, ready for animation, with just a mouse click.

Figure 2: A human participant (right) and the generated avatar (left)

Accessories

To give users a wider sense of likeness, it is possible to incorporate in the generate avatar accessories (Figure 3), like glasses, and change the hair's color. This process is also automated by the scripting solution that we devised for Maya (Figure 4).

Limitations and Future work

The current state of development presents some drawbacks. Although possible to achieve a great level of likeness within a small time frame (around 30 minutes), the model is still not perfect.

Figure 3: A generated avatar with glasses

One of the main drawbacks is clothing. There is still to find a viable solution to mimic the user's clothing in real-time. Not only it is not easy to capture the participant's cloth patterns and shape, but to simulate cloth behavior in real-time is still, nowadays, a complex task, which involves emulating viscous-elastic properties, usually by the means of the finite methods method.

The actual solution is to have all models with a black tight suit, equal to the one worn by users inside the motion capture volume. Once the solution evolves to a body motion capture system without markers, participants will no longer have to wear the black suit, and the clothing will become impairment to the feeling of likeness. Our team is already addressing this issue probing a new technique called Flexistickers (Tzur, 2009), which allows recreating the user's clothing based on pictures.

The next stages of development will be to remove the dependency from external tools – apart from Maya, due to its role in modeling – and provide our own tools for facial acquisition, fitting and marker-less motion capture. The ultimate goal will be to provide a complete pipeline to create a look-a-like avatar in less than 30 minutes, with commodity hardware, giving the possibility to a higher number of people, from field experts, to other scientists conducting experiments aided by virtual reality, to entertainment, to create richer experiences, where users can truly feel immersed in a dream.
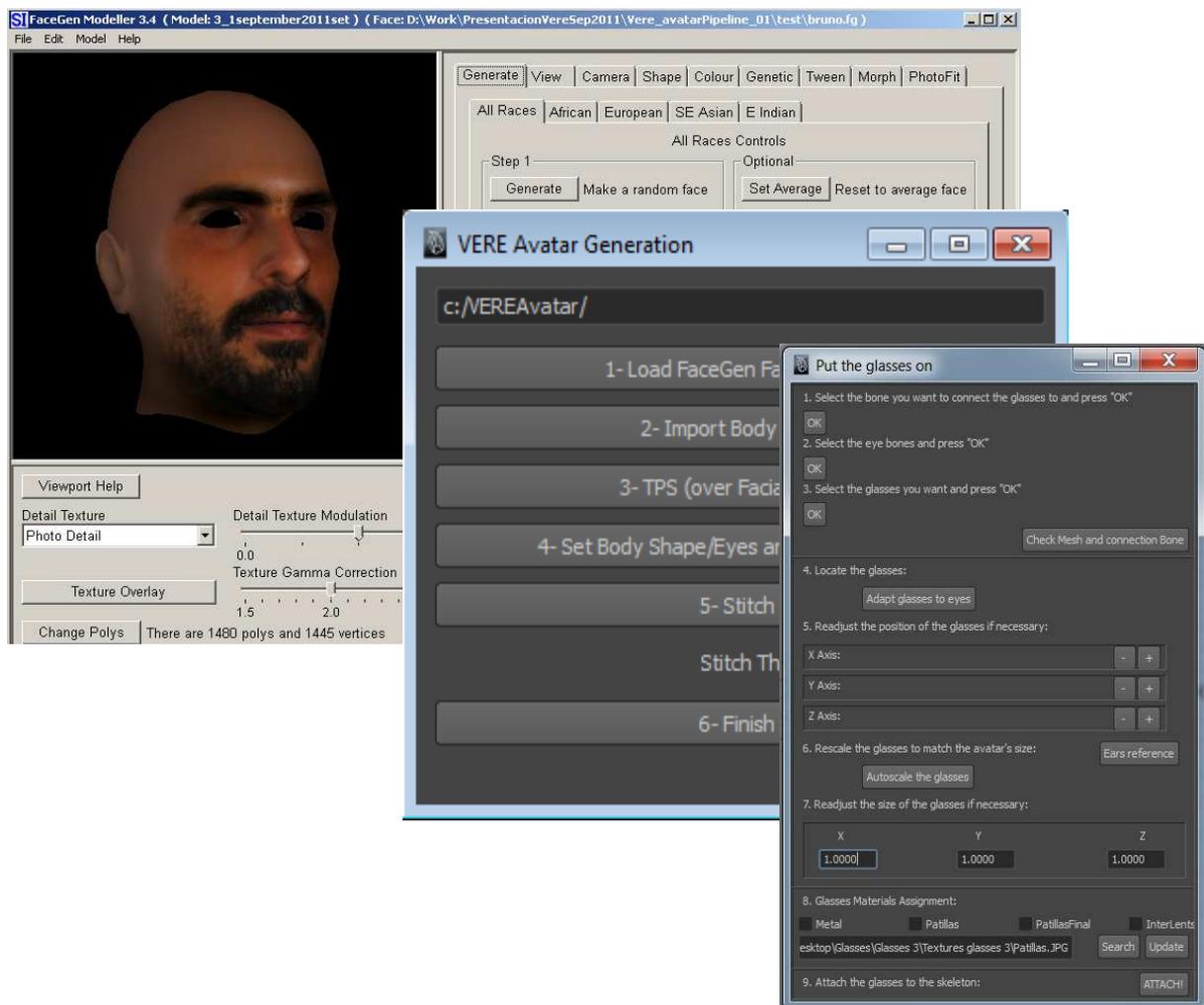
Figure 4: The complete pipeline in tools; FaceGen (Top), Integration UI (middle), Accessories (Bottom)

Bibliography

Autodesk, Inc. (2010). Maya. San Rafael, CA, USA.

NaturalPoint, Inc. (2010). Optitrack. Corvallis, OR, USA.

Orvalho, V. C., Zacur, E., & Susin, A. (2008). Transferring the Rig and Animations from a Character to Different Face Models. *Computer Graphics Forum* , 1997-2012.

Singular Inversions. (2009). FaceGen Customizer. Toronto, Canada.

Singular Inversions. (2009). FaceGen Modeller. Toronto, Canada.

Tzur, Y. a. (2009). FlexiStickers: photogrammetric texture mapping using casual images. *ACM Trans. Graph.* , 45:1--45:10.