A System to Reuse Facial Rigs and Animations

Verónica Costa Orvalho
Instituto de Telecomunicações
Faculdade de Ciências da Universidade do Porto
Porto, Portugal
veronica@faceinmotion.com
http://www.faceinmotion.com

Abstract

Facial animation in films and videogames are strongly dependent on a fixed rig that is custom created for each character. The rig is defined in the early stages of the development and is conditioned by the character's morphology. We present a portable character rigging system that integrates into current animation production pipelines enabling digital artists to create more lifelike characters in less time about 90-99% faster, when compared to traditional animation techniques. It automatically transfers the rig and animations created in one character to different characters, independent of their shape and appearance. Artists are not forced to use predefined rigs and can preserve the original mesh they created. As a result, the system improves the workflow in CG productions, the modeling and animation teams can now work in parallel.

1. Introduction

Facial animation presents many difficulties (time, cost and complexity constraints) that limit its adoption and usefulness in different situations. Pighin et al. [20] discuss the research efforts and main challenges faced by some blockbuster films, and emphasize that facial puppeteering and the use of nonlinear rigs are still unexplored issues. Generating realistic face movements is hard, because even with current 3D software, animators cannot capture and control every detail of the face. To obtain the desired realism, traditional animation pipelines have each character separately rigged by hand, a very labor-intensive and time-consuming task. A rig is a set of controls that allows an artist manipulate a character. The character rigging process is analogous to setting up the strings that control a puppet, which in the hands of an experienced digital artist comes to life [21]. Finding a technique that provides accurate and fast rigging remains a challenge.

Figura 1. Overview: define the source and target model; adapt the source model geometry to fit the target; transfer attributes and shapes; bind the influence objects and skeleton to the target. The result is a model ready to be animated. (Copyright 2005 Dygrafilms)

Our solution overcomes this problem because it adapts the inner structure of the characters to the shape and facial features of the models. Artist can always keep the mesh they created, so the visual look of the characters is never affected. This is fundamental to guarantee the high quality results required by the entertainment industry (e.g. Playable Universal Capture Borshukov et al. [21]), which allows creating and reproducing animations that respect the style and expressiveness of each character, making them unique. Previous methods [18, 7, 27] do not deal directly with the artists needs, and most are oriented towards human look. Our approach is general, so artists can define their own rig and then quickly apply it to different models, even with dis-
parate proportions and appearance (human, cartoon or fantasy). This gives artists complete freedom to manipulate the characters: they can create new animations and not be limited by pregenerated ones. The system we present can easily be integrated into current production pipelines as it is embedded in Maya (Autodesk 2008 [1]). The technology behind the system is described in Orvalho PhD Thesis [19]. It also includes an extensive state of the art analysis related to MPEG-4, FACS another animation techniques. Figure[1] shows an overview of the system.

2. Related Work

Facial animation retargeting between dissimilar meshes is not a new problem, but facial rigging retargeting still unexplored. Most work deals only with the geometry of the face and forgets that the key elements to animate a character is the structure underneath the 3D model mesh. Our work differs from previous approaches that focused on transferring animations, because we aim to transfer the complete facial rig, in addition to animations. We also want to allow the reuse of a facial rig in different face models, regardless of the type of the rig. Thus, most existing work is not efficient for real-time animation, are to complex to setup, force the artist to use a fix pre-defined rigs and there is no previous work capable of dealing with our variety of morphologies [2].

Many efforts have been done to retarget or automatically create the body rig of characters [2,11] and achieved good results, but none have focus on facial rig retargetting. Most facial animation is related to physically-based, geometric deformation and performance-driven methods.

Physically-based methods K. Kahler et al. [14] simulate the contraction and relaxation of human muscles to animate faces. YuanCheng et al. [15] used a multiple-layer dynamic skin and muscle model, together with a spring system, to deform the face surface. But these techniques make it hard to define accurate muscle parameters, due to the complexity of human muscles. So, Sifakis et al. [24] used non-linear finite element implementation to determine accurate muscle action, captured from motion of sparse facial markers. The method shows the success of performance-driven animation, but it is not clear if it can handle anatomically inaccurate models.

Geometric deformation methods use a variety of techniques to animate faces. Following Sederberg and Parry [23]. Chadwick et al. [5] used Free-Form Deformation (FFD) for layered construction of flexible animated characters, which doesn’t require setting the corresponding features on the geometries. Turner and Thalmann [26] used an elastic skin model for character animation. Other approaches were introduced for high level geometric control [16,13] and deformation over 3D model, to help simulate wrinkles [12,25]. These deformation methods provide artists with easy controls to generate animations, but automating these procedures still needs considerable effort.

Performance-driven methods [1] capture the facial performance of an actor, which can be re-targeted to different face models [9,6] or blendshapes [7]. These techniques can generate realistic facial motion, but are expensive to use. Also, they are more suited for human beings than imaginary or fantastic characters.

3. Facial Rigging Challenges

"Rigging is the process of taking a static, inanimate computer model and transforming it into a character that an animator can edit frame-by-frame to create motion" [8]. The result is a rig that can be manipulated by a set of controls like a virtual puppet [22] or by motion capture data. Creating the character rig is a very complex, time consuming and labor intensive task. Still, there is no defined standard methodology for rigging a face. Studios continue to redefine the techniques, processes, technologies and production pipelines to efficiently create films and videogames.

Today, facial animation is done manually by skilled artists, who carefully place and manipulate the animation controls to create the desired motion. As models become more and more complex, it is increasingly difficult to define a consistent rig that can work well for many different characters. So each facial rig has to be created individually by hand. This traditional method ensures high quality results, but it is slow and costly. Large film and videogame companies can afford hiring lots of artists, but this is not feasible for low budget production. It takes an experienced digital artist from one to four weeks to create a complete facial rig, depending on its complexity. But if any change must be applied to an already created rig, the rigging process has to restart. Facial rigging becomes a serious bottleneck in any CG production.

Finding the optimal solution to create a facial rig depends on several constraints: time to develop the rig, budget, artists experience, expected rig performance and actions, and others. The three most common approaches to create a rig are based on: blend shapes [17], bones [28] or a combination of both [16]. However, there are other existing facial animation methods, like motion capture, that can produce photorealistic results and speed up the animation process, but are unable to adapt the performance to dissimilar characters. The captured animation will look the same in all models, ignoring their different appearances. Motion capture focus on analyzing what data to transfer, while our approach focus on what data to transfer and how to represent it.

Thus, the uniqueness of faces makes facial synthesis so challenging. The smallest anomaly in the face shape, proportion, skin texture or movement is immediately detected and classified as incorrect. Most rigging challenges are:
• **no standard**: artists do not follow a formal criteria or methodology when creating a rig, making it difficult to create a solid platform to build upon;

• **changing the geometry or resolution**: it is very common to change the face model during production, to improve the deformation details or simply because it looks better. Any minor modification in the model surface (a bigger nose, more resolution around the lips) after the character is rigged, causes the rigging process to restart;

• **reusing weight maps**: the weight distribution defined for one character will not work on others.

• **number of shapes leads to complex user interface**: many productions use rigs based on hundreds of shapes. Usually, too many shapes make it hard to use the rig. Likewise, if a shape is added during production it can generate two problems: the shape conflicts with existing animations, making it necessary to rework some shots; or the new shape does not mix nicely with the the others;

• **preserving a consistent look**: placing by hand the animation controls leads to different artistic interpretations of where to position each element of the rig. This makes it difficult to easily reproduce the same facial pose between different characters. Consequently, it becomes hard to guarantee a consistent look throughout the production;

4. **Our Facial Rigging System**

Creating and placing by hand each component of the rig (bones, controls) quickly becomes impractical when complexity grows. The system we present can handle simple and complex rigs based on a new approach described in our previous work [19]. The system is:

• **generic**: the facial rig can have any type of configuration and does not force the use of a predefined rig;

• **flexible**: the rig has no initial constraints;

• **independent of the shape**: a facial rig can be transferred between models that have different geometry, look and appearance;

• **enhances artistic freedom**: artists can use any tool or deformation method to create the rig.

![Figure 2. The first row shows the source model (Lisandro) and the rest show the target models (Mostaza, Teseo, Hada and Demetrio); first column shows the look and appearance of the models; second column details the facial rig that includes 21 joints and 1 NURBS surface; and third column shows the weight distribution. All models have different wireframe. (Copyright 2005 Dygrafilms)](image-url)
and can be created by an artist. The second model, we call target model, doesn’t have a character rig associated to it. The source and target models can have different descriptors: one can be defined as a polygonal surface and the other as a NURBS surface. Also, the faces do not need to have the same number of vertices. Figure 1 shows an overview of the system pipeline and illustrates the rig transfer process with two dissimilar characters. The main steps within the system are: 1. surface deformation; 2. attribute transfer; 3. skinning.

1. Surface Deformation The source rig information is used as the direct input for transferring the setup to the target model. First, our deformation method deforms the source model surface to match the geometry of the target. We landmark the facial features to keep correspondence between source and target model, and then employ a computer vision interpolation technique named Thin Plate Splines (TPS) [4], as our deformation kernel function. After the TPS, the source surface only has exact deformation at the landmark positions of the target model, while the rest of the points lay outside the target surface. We solve this by applying a dense correspondence algorithm [19], which projects every point of the warped surface to the closest point of the target and determines the correspondence between every source and target vertex.

2. Attribute Transfer Using as reference the previously deformed source surface, we call guide model, the method accurately places the source rig attributes (section 4.2 describes the rig attributes) into the target model, even if they have different geometric proportions. We had to adapt the TPS to properly deal with each attribute specific characteristics. It is not the same transferring bones than transferring a NURBS curve. The dense correspondence avoids placing additional landmarks on the influence objects or on the skeleton structure. The deformation process achieves excellent results in positioning the source rig attributes in the correct regions of the target face. For example, joints and NURBS surfaces are relocated in the target model, based on the correspondent position they have in the source model. They are also transformed to fit the shape and size of the target (see figure 2).

3. Skinning After the deformation step comes the skinning, based on a smooth binding algorithm. It binds the transferred attributes to the target model using the adjusted weights of the source, avoiding the need for manual weighting. The weights at the target are calculated using the deformation method. Each vertex of the target model accurately adapts the blending weight of the joints and influence object, based on the source model weight distribution, to properly represent the target facial look and behavior (see figure 2). Last, as the target model is already rigged and weighted, transferring facial animations is a straightforward process. The method only needs to scale and adapt the animation curves to fit the proportions of the target. The end result are face models ready to be animated with production quality rigs.

4.2. Rig Definition

Central to our system is the notion of source rig $S$, and we use the model in figure 2 to illustrate it. The rig is formed by different layers of abstraction that we refer to as attributes: skin surface $S_S$, influence objects $S_O$, skeleton bones $S_B$, facial features landmarks $\lambda$, shapes $S_H$, animation scripts $S_A$ and other components for representing the eyes, teeth and tongue. We can assign additional attributes to each of these layers: weight, texture, etc. [10].

The source rig helps define the appearance of the characters. It establishes the character setup standard shared by all the models. Artists can create their own source rig, because they are free to use any type of controls and components to achieve the desired visual look.

The source rig $S$ has been modeled manually and is a highly deformable structure of a face. During the modeling process, we used facial features and regions to guarantee realistic animation and reduce artifacts.

The surface $S_S$ is the external geometry of the character that determines the skin of the face, using polygonal surfaces composed by a set of vertices $r$ and a topology that connects them.

The source rig is tagged with landmarks $\lambda$, distributed as a set of sparse anthropometric points. We use the landmarks to define specific facial features to guarantee correspondence between models.

The skeleton $S_B$ is a group of bones positioned under
The skin. It defines the pose of the head and controls lower level surface deformation. Each bone is defined by two joints, one at each end of the bone.

The influence objects $S_I$ are objects that affect the shape of the skin and help artists control the 3D models. They include: NURBS surfaces, NURBS curves, lattice deformers, cluster deformers, polygon mesh, and others.

The shapes $S_H$ are new 3D face models created by applying deformations over the geometry $S_S$ of the character. A shape is a 3D facial pose of the source model, where $S_H$ and $S_S$ have the same geometry. Shapes are usually modeled manually by an artist. They represent facial expressions or partial deformation of a specific area of the face. They are used to create blend shapes, which let you change the shape of one object into the shapes of other objects. The interpolation between shapes results in facial animations.

The animation scripts $S_A$ consist of a list of animation curves that determine motion. Each animation curve represents changes in the value of an attribute, like shapes or bones.

4.3. Application and Workflow

We implemented a set of plug-ins in C++ for Maya. The plug-in includes a simple user interface to ease the landmarking and assist the transfer process (see figure 4). The modular design of the application makes it simple to integrate into existing animation pipelines.

The application enables artists to fit automatically the rig from the source to the target model; manipulate the target as if they were using a puppet; and adjust animation parameters in the target model or animate the target using predefined source animations.

Figura 4. Application user interface running on Maya: assisting the landmarking process (left); close up of the source rig viewer with joints and influence object selected (right). (Models copyright 2005 Dygrafilms)

The input to the pipeline is the source model $S$ information. The output is a fully rigged target model $F$ ready to be animated. The workflow of the application is as follows:

1. Landmarking: Defines the source and target model landmarks that will keep correspondence between models.

2. Surface Correspondence: Ensures the exact point matching at the landmarks and smoothly interpolates the deformation of other points.

3. Surface Dense Correspondence: Ensuring exact deformation of every surface point, avoids placing additional landmarks.

4. Attribute Transfer: Uses the TPS deformation method to transfer each type of attribute.

5. Skinning: Binds the deformable objects, influence objects and surface to the skeleton of the target model.

Figura 5. System pipeline. Shows the three main steps needed to transfer a facial rig: skin deformation, attribute transfer and skinning. The output of the skin deformation is the guide model, which serves as reference for the rest of the transfer process. The output of the attribute transfer is the target model with the rig components positioned in correspondence to the source. Last, after skinning the character using the source model weights, the target model is rigged and ready to be animated. (Copyright 2005 Dygrafilms)
5. Results and Discussion

Reproducing the subtleties of a face through animation requires developing a sophisticated character rig. But, creating by hand the inner structure and controls of each character is a very labor-intensive and time-consuming task. We presented a system that transfers the rig and animations between characters, at least an order of magnitude faster than traditional manual rigging. The system allows creating the rig, animation controls and scripts for one model (source), and reuse them in many different target models. It is independent of the appearance or shape of the model, so rig transfer between dissimilar characters is feasible. Artists can create their own rigs and are not forced to use predefined ones. In film and videogame productions, artists are often given one base model to make all new faces (shapes). Also, it is common that afterwards they are asked to use a different 3D face, because it has improved deformation details or simply looks better. Currently, all shapes need to be remade to reflect the topology of the new face. But our method makes sure that previous work can be transferred and artists time is not wasted.

Our facial animation system can be integrated into existing animation production pipelines, improving its work flow as it decouples the work of the animators and the modelers, they can be working in parallel on the same character. As a result, the companies will have fewer bottlenecks, which will increase productivity and reduce cost.

The system also provides a solid foundation for setting up consistent rigging strategies: at the beginning of a production, artists can define the required rig parameters and afterwards use them as a template for all models. This rig becomes the building block for all characters. Our approach helps film and videogame studios overcome the current lack of a standard rigging methodology. It guarantees that all rigs generated by the system produce homogeneous results, ensuring that the models share a common vision and consistent artistic style.

Testing and Validation We validated the system with a series of experiments. We used source models from several companies (Electronic Arts, Blur Studios, Dygrafilms, etc.) and for each, we transferred the rig and animations to different target models. We worked with a variety of styles: human, cartoon and fantastic creatures. Then, for the same models, we compared the output of our application with the results manually created by an artist. The results were supervised by Technical and Art Directors, who approved the quality of our rig and animations to be used in CG productions, replacing the artist generated ones (see figure 6, for a detail explanation). This is a crucial result: if the output still requires a lot of tuning, then the system is useless in a production. The examples of the paper are limited to synthetic characters to emphasize the versatility of the method.

Performance Our application allows creating the rig in one hour as we need to visually validate the results, (go through all the rig) which take some time if the rig is very complex. The attribute transfer process like changing the weights, modifying a control position or transferring animations, is nearly instantaneous. Figure 7 shows that our method convincingly captures the complex effect of simulating a talking head, to be used in a film. The big time savings achieved on the rigging process is usually an order of magnitude or more, and still meet the high quality animation needs of the entertainment industry. The tests were made on a AMD Athlon 64 3500+ CPU with 2 GB of RAM.

Extreme Test To test the method to the extreme of its
possibilities, we ordered three very different 3D models: a photorealistic human (source model), a cartoon and a fantastic creature (target models). The models differ enormously in artistic style, deformation behavior, shape and proportions (see figure 8). The source model rig includes: 2 NURBS curves around the mouth, 1 lattice for the jaw, 6 joints for the head, 5 joints for the tongue, 3 joints for the teeth, 47 shapes and 2 animations clips (one with extreme facial poses and the other with lip-sync). The method successfully transferred the shapes in the mouth region, which is very complex due to the variety of poses it can perform. But the drawback of transferring shapes between characters with different styles is that the target models inherit the movements of the source. We obtained a cartoon model simulating a human character. This faithfulness it is not always what the artist wants, so we needed to keep in mind this behavior.

Figura 8. Source (top) and target models (bottom). Keyframes extracted from a video sequence to show different poses transferred from the source to the target models. The poses were created by manipulating the source rig. (Copyright 2007 Face In Motion)

Limitations An important issue to mention is that if the source rig quality is low, the transference is still successful but the results on the target will be of comparable quality. The technology is indeed independent of the quality and the shape of the rig. During our tests, we realized that when the source model has the eyes and the mouth completely closed, the attribute transfer results show some artifacts. This a current limitation of our solution. To obtain artifact free transfers, it is recommended to have the eyes and mouth of the source and target models slightly opened.

Future Work We performed some tests on mapping motion capture data into the source rig, and later transfer it to the target model. This is an interesting direction for future research and to extend our application.

6. Acknowledgement

Many thanks to Juan Nouche (Ottiplanet) and Xenko Alvarez (Enne Studios) for their feedback and for providing the 3D models while working at Dygrafilms. Also to Fred Fowels (Rainmaker), Jean Luc-Duprat (Intel) and Crystal Wang (Electronic Arts) for testing the system and providing 3D models and motion capture data. Special thanks to Toni Susin for supervising the research project. Last, we would like to thank Blur Studios for using our system into their CG pipeline to create the facial rig of the Simpsons and Fable productions. This project is partially funded by FCT, Portugal (http://www.it.pt).

Referências


