Transferring a Labeled Generic Rig to Animate Face Models

Verónica Costa Teixeira Orvalho[1], Ernesto Zacur[2], and Antonio Susin[1]

[1] Laboratorio de Simulación Dinámica (Univ. Politècnica de Catalunya)
veronica.costa@upc.edu, toni.susin@upc.edu
[2] Universitat Pompeu Fabra
ernesto.zacur@upf.edu

Abstract. We present a facial deformation system that adapts a generic facial rig into different face models. The deformation is based on labels and allows transferring specific facial features between the generic rig and face models. High quality physics-based animation is achieved by combining different deformation methods with our labeling system, which adapts muscles and skeletons from a generic rig to individual face models. We describe how to find the correspondence of the main attributes of the generic rig, transfer them to different 3D face models and generate a sophisticated facial rig based on human anatomy. We show how to apply the same deformation parameters to different face models and obtain unique expressions. Our goal is to ease the character setup process and provide digital artists with a tool that allows manipulating models as if they were using a puppet. We end with different examples that show the strength of our proposal.

1 Introduction

Facial animation is related to the interaction of muscles and skeletons beneath the skin. It is the key element to transmit individuality and personality to a character in films and video games. Therefore, to obtain physically-based animations, it is crucial to develop systems that simulate the anatomical structure of the face. Recent advances in facial synthesis show an increased interest in physics-based approaches [23] [15] [22]. Today, to animate a character, an experienced CG artist has to model each facial rig by hand, making it impossible to re-use the same rig in different facial models. The task is further complicated when a minor artistic change on the facial topology leads to the restarting of the rigging process from scratch. This creates a bottleneck in any CG production and leads to the research of automated methods to accelerate the process [14].

Modeling and animation of deformable objects have been applied to different fields [1] [3]. *Noh et al.* [17] proposed several methods for transferring animations between different face models. The surface correspondence is obtained by specifying the corresponding point pairs on the models. *Pighin et al.* [7] presented a method to interactively mark corresponding facial features in several photographs of an individual, to deform a generic face model using radial basis function. *Sederberg and Parry* [20] first introduced Free-Form Deformation

F.J. Perales and R.B. Fisher (Eds.): AMDO 2006, LNCS 4069, pp. 223–233, 2006.

(FFD) in 1986; the method does not require setting the corresponding features on the geometries. Other interesting approaches for high level geometric control and deformation over 3D model were introduced [5] [12] [21].

We propose a deformation method to transfer the inner structure of a generic rig to individual face models, based on thin-plate splines [2] and the use of facial features labels. We tag the generic rig with landmarks on its surface (the skin) and automatically deform it, together with the muscle and skeleton structure, to fit different face models. Because all models share the same generic set of attributes, we don't need to develop unique scripts for each face. We can transfer generic rig parameters, enabling re-use of existing animation scripts. We can build models with underlying anatomical structure, skin, muscle and skeleton, for human heads or other type of creatures. The models are suitable for real-time animation based on simulation of facial anatomy.

2 The Generic Rig

Our method builds on a sophisticated 3D face model we call generic rig $\mathcal{R}$ (see figure 4), designed for use within a facial animation production pipeline to accelerate the rigging process. The model is formed by different layers of abstraction: skin surface $\mathcal{R}_S$, muscles surfaces $\mathcal{R}_M$, skeleton joints $\mathcal{R}_B$, facial feature landmarks λ, skinning system and other components for representing the eyes, teeth and tongue. We can assign different attributes to each of these layers, like: weight, texture, muscle stress, etc. [10]

The **generic rig** $\mathcal{R}$ has been modeled manually and is a highly deformable structure of a face model based on physical anatomy. During the modeling process, we used facial features and regions to guarantee realistic animation and reduce artifacts.

The **surface** $\mathcal{R}_S$ is the external geometry of the character, determining the skin of the face using polygonal surfaces composed by a set of vertices $\mathbf{r}$ and a topology that connects them.

The generic rig is tagged with **landmarks** λ, distributed as a set of sparse anthropometric points. We use these landmarks to define specific facial features to guarantee correspondence between models. Our rig has 44 landmarks placed on the surface (see figure 4c) [9] [6].

The **skeleton** $\mathcal{R}_B$ is a group of bones positioned under the skin. It defines the pose of the head and controls lower level surface deformation.

The **muscles** $\mathcal{R}_M$ are a group of volumes, surfaces or curves located under the skin, which control higher level surface deformation. To build our muscle structure, we selected eleven key muscles (see figure 4d) responsible for facial expressions [8], out of the twenty-six that move the face.

3 Transferring the Generic Rig Structure

We introduce a method to automatically transfer the generic rig structure and components to individual 3D face models, which can be divided in three main steps: first, we deform the generic rig surface to match the topology of the face model we want to control; then, we adapt the muscles, skeleton and attributes of the generic rig to the 3D model; finally, we bind the transferred elements to the model, obtaining an anatomic structure prepared for physically-based animation.

The face model that inherits the generic rig setup is referred as $\mathcal{F}$. It is defined by a face surface $\mathcal{F}_S$, which determines the face geometry and shape, and a set of landmarks ϕ placed on $\mathcal{F}_S$. Like $\mathcal{R}_S$ from the generic rig, $\mathcal{F}_S$ is defined by a set of vertices $\mathbf{f}$ and a topology that connects them. The landmarks are positioned manually by the artist, to guarantee correspondence with the generic rig landmarks (see section 2). Even though the generic rig has 44 landmarks, it is not necessary to use them all to transfer the rig (see results in figure 5). Starting with a landmarked face model $\mathcal{F}$, the rest of the structure transfer is automated as it will be detailed next.

3.1 Geometric Transformations

To deform the rig $\mathcal{R}$ into $\mathcal{F}$ we use linear and non-linear global transformations and local deformation. Linear transformations in combination with non-linear transformations, give us enough degrees of freedom (DOF) to ensure the correct match between the geometries.

Equation 1 describes the generic form of the transformations:

$$\mathbf{x}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} w_{xi} U(\mathbf{x}, \mathbf{p}_i) + a_{x0} + a_{xx}\, x + a_{xy}\, y + a_{xz}\, z \\ \sum_{i=1}^{n} w_{yi} U(\mathbf{x}, \mathbf{p}_i) + a_{y0} + a_{yx}\, x + a_{yy}\, y + a_{yz}\, z \\ \sum_{i=1}^{n} w_{zi} U(\mathbf{x}, \mathbf{p}_i) + a_{z0} + a_{zx}\, x + a_{zy}\, y + a_{zz}\, z \end{pmatrix} \tag{1}$$

Following Bookstein [2] [18], we use the kernel function $U(\mathbf{x}, \mathbf{p}_i) = \|\mathbf{x} - \mathbf{p}_i\|$ that minimizes the bending energy of the deformation. This transformation is called Thin Plate Spline Warping (TPS) and it is a special case of Radial Basis Function Warping [4].

Solving the linear system of equations 2, we obtain $\mathbf{w}$ and $\mathbf{a}$ coefficients, using $\mathbf{p}$ and $\mathbf{q}$ correspondence, where $\mathbf{p}$ are surface origin coordinates and $\mathbf{q}$ are surface target coordinates. The TPS wrapping ensures the exact point matching and interpolates the deformation of other points smoothly.

$$\left(\begin{array}{cccc|cccc} 0 & U(\mathbf{p}_1,\mathbf{p}_2) & \cdots & U(\mathbf{p}_1,\mathbf{p}_n) & p_{x1} & p_{y1} & p_{z1} & 1 \\ U(\mathbf{p}_2,\mathbf{p}_1) & 0 & \cdots & U(\mathbf{p}_2,\mathbf{p}_n) & p_{x2} & p_{y2} & p_{z2} & 1 \\ \vdots & & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ U(\mathbf{p}_n,\mathbf{p}_1) & U(\mathbf{p}_n,\mathbf{p}_2) & \cdots & 0 & p_{xn} & p_{yn} & p_{zn} & 1 \\ \hline p_{x1} & p_{x2} & \cdots & p_{xn} & 0 & 0 & 0 & 0 \\ p_{y1} & p_{y2} & \cdots & p_{yn} & 0 & 0 & 0 & 0 \\ p_{z1} & p_{z2} & \cdots & p_{zn} & 0 & 0 & 0 & 0 \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 & 0 \end{array}\right) \left(\begin{array}{c} w_1 \\ w_2 \\ \vdots \\ w_n \\ \hline a_x \\ a_x \\ a_x \\ a_0 \end{array}\right) = \left(\begin{array}{c} q_1 \\ q_2 \\ \vdots \\ q_n \\ \hline 0 \\ 0 \\ 0 \\ 0 \end{array}\right) \tag{2}$$

3.2 Surface Deformation

Given $\mathbf{p}$ and $\mathbf{q}$, we define the operation:

$$\mathbf{x}' = TPS_{\mathbf{p}}^{\mathbf{q}}(\mathbf{x}) \tag{3}$$

that minimizes the energy of the surface deformation. We use the following notation, $\mathbf{q} = \mathbf{p}|_S$, where $\mathbf{q}_i$ is the position of the correspondent point to $\mathbf{p}_i$ in the geometry S.

Figure 1a shows the deformation of a surface uniformly sampled into another surface, using a reduced set of sparse landmarks. Only these landmarks will result on an exact deformation, while the rest of the surface points lay outside the target surface. Figure 2 shows the deformation of the generic rig into a face model using 10 anthropometric landmarks.

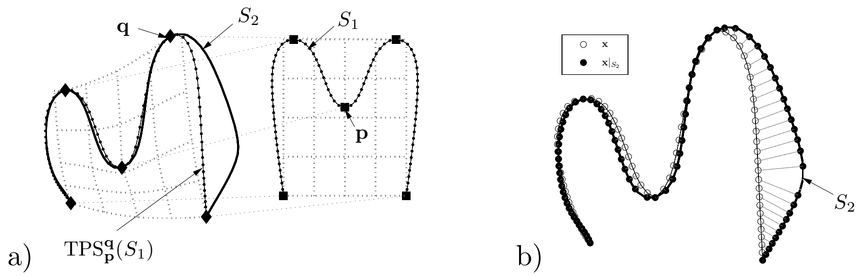


Fig. 1. a) TPS wrap of a generic surface based on reduced set of sparse landmarks (S_1: original surface, S_2: target surface, **p**: origin landmarks, **q**: target landmarks); b) Sticking of the original surface to the target surface after applying the TPS (see section 3.3)

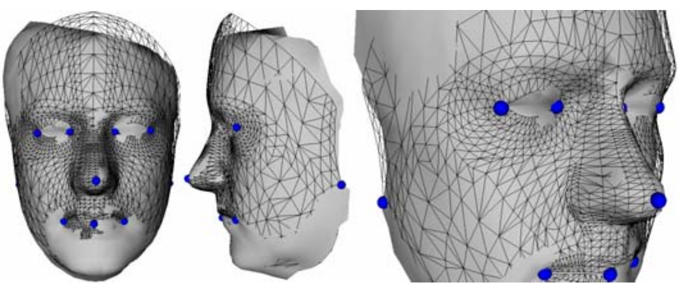


Fig. 2. Human Face wraping process using 10 landmarks

3.3 Obtaining a Dense Correspondence Between Surfaces

To obtain an exact deformation of every surface point, where the origin surface matches the target surface, we apply a local deformation to every point of the origin surface. Then, we project every point of the wrapped surface to the closest point of the target surface. As a result, we get the correspondent point in the target surface for every vertex of the origin surface. This is called *dense correspondence* [13] between surfaces.

We define in our pipeline an operation called *Stick* (STK) that computes the dense correspondence of points $\mathbf{r}$, between the generic rig $\mathcal{R}$ and the face model $\mathcal{F}$:

$$\mathbf{r}|_{\mathcal{F}} = STK_{\mathcal{F}_S}\left(TPS_{\lambda}^{\phi}(\mathbf{r})\right) \tag{4}$$

This operation can present undesirable folds in areas with high curvature or if the distance between origin and target points is large. Lorenz and Hilger worked on solutions to avoid these folds [16] [11]. Fortunately, we didn't came across this problem in the many tests we performed on different face models: human and cartoon.

3.4 Deforming Layer Structures

Based on the dense correspondence between $\mathcal{R}_S$ and $\mathcal{F}_S$, we can deform the generic rig muscles $\mathcal{R}_M$ and skeleton $\mathcal{R}_B$. This correspondence avoids placing additional landmarks on the muscles or on the skeleton structure. Figure 3 shows that the wrap based on dense correspondence keeps the relationship between the structure and the surfaces better than the wrap based on sparse landmarks.

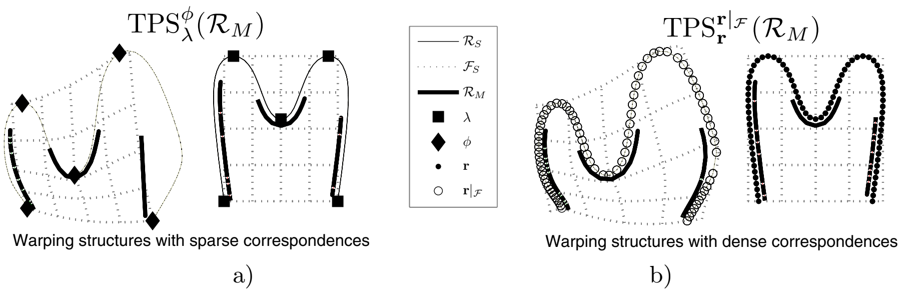


Fig. 3. Wrap based on a) landmarks; b) dense correspondence

3.5 Attribute Transfer

The generic rig $\mathcal{R}$ has a set of attributes on the surface nodes $\mathbf{r}$ defined as scalar or vectorial fields. We have to transfer each of these attributes to surface $\mathcal{F}_S$. For each surface vertex $\mathbf{f}_i$, we find its closest point on $\mathcal{R}_S|_{\mathcal{F}}$, get the interpolated value and assign it to $\mathbf{f}_i$.

Figure 9a shows the transferred weights that influence the movement of the jaw bone. Figure 9b shows a region labeling transfer. Both figures show the attributes transfer from the generic rig to the cartoon, with different triangulations.

3.6 Skinning

In animation, skinning is the process of binding deformable objects to a skeleton [19]. In some software packages it is also known as envelope or birail. After skinning, the deformable object that makes up the surface is called the character's skin, and the deformable objects under the skin, which influence and shape it, are called the muscles.

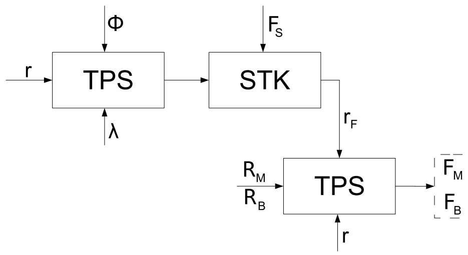
The output of the skinning process is a character model setup, with the skeleton and muscles controlling the deformations. The positioning of the muscles has two goals: build an inner structure that correctly reflects the character's appearance and enable the projected facial animations with minimum effort. The deformations of the character's skin, produced by the movements of the skeleton and muscles, allows physically-based animation.

Our skinning method uses the generic rig weight to automatically attach the previously deformed skeleton and muscles to the face model $\mathcal{F}$.

3.7 Method Overview

Next, we describe the method pipeline:

$\mathcal{R} \leftarrow$ Generic Rig
$\mathcal{F} \leftarrow$ Face Model
$\lambda \leftarrow$ Generic Rig Landmarks
$\phi \leftarrow$ Face Model Landmarks



$\mathcal{R}'_S \leftarrow \mathrm{TPS}^{\phi}_{\lambda}(\mathcal{R}_S)$
$\mathbf{r}|_{\mathcal{F}} \leftarrow \mathrm{STK}_{\mathcal{F}}(\mathcal{R}'_S)$
$\mathcal{F}_M \leftarrow \mathrm{TPS}^{\mathbf{r}|_{\mathcal{F}}}_{\mathbf{r}}(\mathcal{R}_M)$
$\mathcal{F}_B \leftarrow \mathrm{TPS}^{\mathbf{r}|_{\mathcal{F}}}_{\mathbf{r}}(\mathcal{R}_B)$
$\mathbf{f} \leftarrow attributeTransfer(\mathbf{r}|_{\mathcal{F}})$
$\mathcal{F} \leftarrow skinning(\mathcal{F}_S, \mathcal{F}_M, \mathcal{F}_B)$

4 Results and Conclusion

The deformation methods have been implemented in C++ as a plug-in for Maya 7.0 software. Our method speeds up the character setup and animation pipeline, since we drive all face models by deformation of the same generic rig. This allows using the facial expressions created in the rig on different models. To obtain unique deformation in each face, both generic rig's muscles and skeleton can be adjusted in the different facial regions.

In contrast with other methods [15] that landmark the skin, muscle and skull, we only landmark the skin surface because we obtain dense correspondence. This simplifies and eases the setup of the character. Our results indicate that anthropometric modeling is a good approach to generate physically-based animations.

Our *generic rig* has 1800 points, 44 landmarks, 4 bones and 11 muscles, and is based on human anatomy (see figure 4). The *human model* is a 3D scan of a human face. It has 1260 points and 10 landmarks (see figure 5). Figure 5b displays the wireframe mesh. We use 10 landmarks to transfer the rig structure (see figure 5d). Figure 12 shows the wrapping process.

The *cartoon model* has 1550 points and 44 landmarks (see figure 6). Figure 6 shows the muscle transfer and figure 9 shows the attribute transfer of the weight and region label. Based on the weights of figure 9a, figure 10 shows the transfer of a facial expression. The graphics on figure 8 display the distance between the muscle and the skin surface points, on the generic rig (solid line) and on the face model (dots). Results show that the wrapping works better for human faces. To explore the limits of our method, figure 7 confirms that the wrapping and landmarks fitting work robustly in non-human faces with extreme facial appearance. We use 12 landmarks to transfer the rig structure to a goat (see figure 7d).

For further automation we will create a set of facial expression templates and an intuitive GUI running in Maya. Our generic rig will include different type of muscles. We will add support on our plug-in for NURBS surfaces. We will allow the models to inherit the animation controls from the generic rig. The purpose of these animation controls is to reduce: the complexity to obtain facial motion, the effort required by artist and computation time.

Our final goal is to automate the character setup process within an animation pipeline, without changing the input model, enabling the artists to manipulate it as if they were using a puppet. The model can be created by an artist or scan generated. This will further speed up the creation of animations, because it will require no additional rigging.

Acknowledgement

Special thanks goes to João Orvalho for his review, unconditional support and motivation. We also thank Dani Fornaguera, Marco Romeo and Carlos for their valuable comments and 3D Models. This research is partially supported by CI-CYT grant TIN2004-08065-C02-01.

References

[1] A. Angelidis, M. Cani, G. Wyvill, and S. King, *Swirling-sweepers: Constant-volume modeling*, Pacific Graphics 2004, 2004.

[2] F. Bookstein, *Principal warps: Thin-plate splines and the decomposition of deformations*, IEEE Trans. on Pattern Anaylsis and Machine Intelligence, vol. 11, no. 6, 1989, pp. 567–585.

[3] M. Botsch and L. Kobbelt, *An intuitive framework for real-time freeform modeling*, ACM Transactions on Graphics (TOG), SIGGRAPH '04, 2004, pp. 23(3), 630–634.

[4] J. Carr, W. Fright, and R. Beatson, *Surface interpolation with radial basis functions for medical imaging*, vol. 16, IEEE Trans. on Medical Imaging, 1997.

[5] S. Coquillart, *Extended free-form deformations: A sculpturing tool for 3d geometric modeling*, Proc. SIGGRAPH 90' Conf., ACM Computer Graphics, 1990, pp. 187–196.

[6] D. Metaxas D. DeCarlo and Matthew Stone, *An anthropometric face model using variational techniques*, Proc. SIGGRAPH '98, 1987, pp. 67–74.

[7] R. Szeliski D.H. Salesin F. Pighin, D. Lischinski and J.Hecker, *Synthesizing realistic facial expressions from photographs*, Proc. SIGGRAPH '98 Conf, 1998, pp. 75–84.

[8] G. Faigin, *The artist's complete guide to facial expressions*, Watson-Guptill Publications, New York, 1987, pp. 67–74.

[9] Munro Farkas, Leslie and Ian, *Anthropometric facial proportions in medicine*, Charles Thomas publisher ltd., USA, 1987.

[10] J. Haber, *Anatomy of the human head*, SIGGRAPH 2004, Course Notes: Facial Modeling and Animation, 2004.

[11] K. B. Hilger, R. R. Paulsen, and R. Larsen, *Markov random field restoration of point correspondences for active shape modelling*, SPIE - Medical Imaging, 2004.

[12] W.M. Hsu, J.F. Hugues, and H. Kaufman, *Direct manipulation of free-form deformation*, Proc. SIGGRAPH '92, ACM Press, 1992, pp. 177–184.

[13] T. Hutton, B. Buxton, and P. Hammond, *Dense surface point distribution models of the human face*, IEEE Workshop on Mathematical Methods in Biomedical Image Analysis, 2001, pp. 153–160.

[14] P. Joshi, W. Tien, M. Desbrun, and F. Pighin, *Learning controls for blend shape based realistic facial animation*, Eurographics/SIGGRAPH Symposium on Computer Animation, ACM Press, 2003, pp. 187–192.

[15] H. Yamauchi H. Seidel K. Kahler, J. Haber, *Head shop: Generating animated head models with anatomical structure*, ACM, 2002.

[16] C. Lorenz and N. Krahnstöver, *Generation of point-based 3d statistical shape models for anatomical objects*, vol. 77, Computer Vision and Image Understanding: CVIU, 2000, pp. 175–191.

[17] J. Noh and U. Neumann, *Expression cloning*, Proc. SIGGRAPH '01 Conf, ACM SIGGRAPH, 2001, pp. 277–288.

[18] K. Rohr, H.S. Stiehl, R. Sprengel, T.M. Buzug, J. Weese, and M.H. Kuhn, *Landmark-based elastic registration using approximating thin-plate splines*, vol. 20, IEEE Trans. on Medical Imaging, 2001, pp. 526–534.

[19] J. Schleifer, *Character setup from rig mechanics to skin deformations: A practical approach*, Proc. SIGGRAPH '02, Course Note, 2002.

[20] T. Sederberg and S. Parry, *Free-form deformation of solid geometric models*, Proc. SIGGRAPH 86' Conf., ACM Computer Graphics, 1986, pp. 151–160.

[21] K. Singh and E. L. Fiume, *Wires: a geometric deformation technique*, Proc. SIGGRAPH 98' Conf., ACM Computer Graphics, 1998, pp. 405–414.

[22] R. Szeliski and S. Lavallee, *Matching 3d anatomical surfaces with non-rigid deformation using octree splines*, Internatinal Journal of Computer Vision 18,2, 1996, pp. 171–186.

[23] K. Waters and J. Frisbie, *A coordinated muscle model for speech animation*, Proc. Graphics Interface '95, 1995, pp. 163–170.

Transferring a Labeled Generic Rig to Animate Face Models

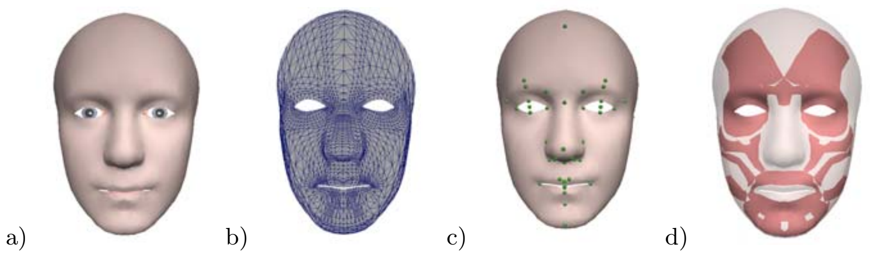


Fig. 4. Generic Rig a)textured; b)wireframe; c)44 landmarks; d)muscles

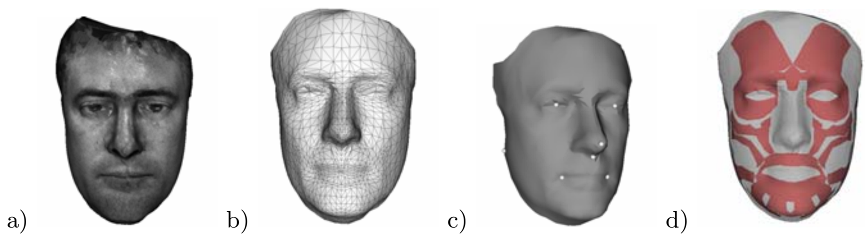


Fig. 5. Human a)textured; b)wireframe; c)10 landmarks; d)muscles

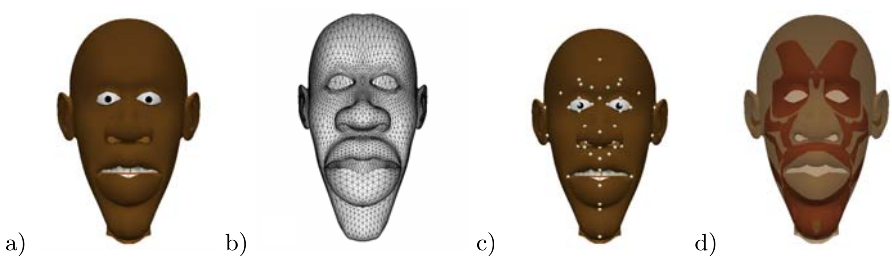


Fig. 6. Cartoon a)textured; b)wireframe; c)44 landmarks; d)muscles

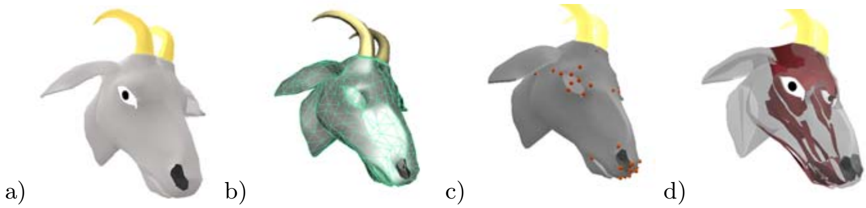


Fig. 7. Animal a)textured; b)wireframe; c)44 landmarks; d)muscles

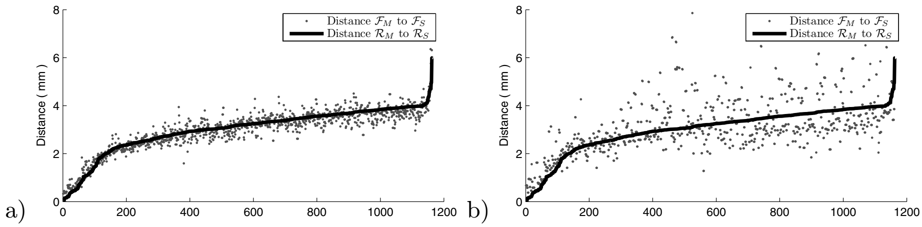


Fig. 8. Distance between muscle and skin surfaces on the generic rig and on the model a) Human model; b) Cartoon model

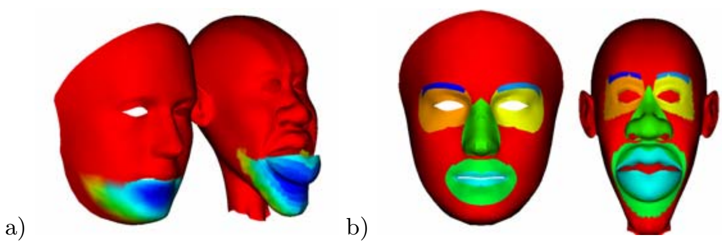


Fig. 9. Attribute transfer from generic rig to cartoon model a) weight of the jaw bone (red is $w = 0$, blue is $w = 1$); b) region labels

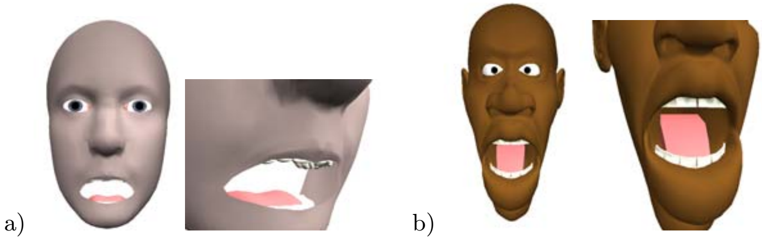


Fig. 10. Facial Expression a) Generic Rig and close up; b) Cartoon and close up

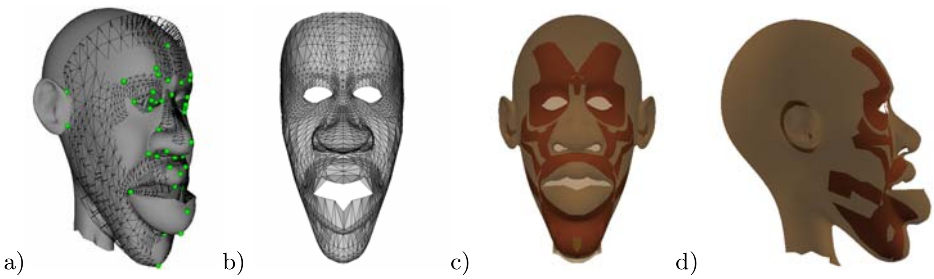


Fig. 11. Cartoon Deformation a) TPS and Stick Lines; b) Cartoon after STK; c) Muscles transfer front view; d) Muscles transfer side view

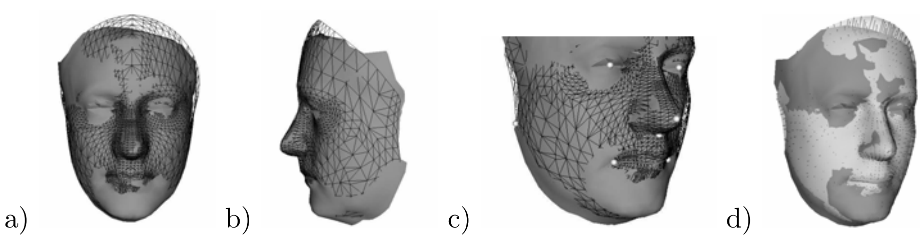


Fig. 12. Human Face Deformation with 10 landmakrs a) TPS front view; b) TPS side view; c) close up; d) dense correspondence after STK