A Physical Model of Facial Tissue and Muscle Articulation

Keith Waters† and Demetri Terzopoulos††

†Schlumberger Laboratory for Computer Science, P.O. Box 209015, Austin, TX 78720
†Department of Computer Science, University of Toronto, Toronto, Ontario, M5S 1A4

Abstract

This paper presents a physically-based facial tissue model of the epidermis, subcutaneous fatty tissue, and skeletal facial foundation. The tissue model is represented by a lattice of non-linear spring units with varying visco-elastic properties. Muscle force vectors are constructed from fibers that are rigidly attached to the bone at one end and blended into the dermis layer at the other. A numerical integration technique is employed to calculate the dynamic displacement of nodes in the lattice under the influence of internal muscular contractions. An expressive 3D face generated on a high performance graphic workstation demonstrates the facial tissue and muscle articulation model.

1 Introduction

The abiding interest in the human face is derived from its shape and remarkable powers of expression. Facial expressions can convey intricate signals to other humans making the face an extremely powerful form of non-verbal communication. This paper develops a modular approach to biomimetic facial animation. Muscle force vectors are constructed from fibers that are rigidly attached to the bone at one end and blended into the dermis layer at the other. A numerical integration technique is employed to calculate the dynamic displacement of nodes in the lattice under the influence of internal muscular contractions. An expressive 3D face generated on a high performance graphic workstation demonstrates the facial tissue and muscle articulation model.

1.1 Overview

Section 2 places our work in perspective before the background of prior investigations into facial modeling and animation. After a brief review of the structure and mechanical properties of facial tissue, Section 3 describes our tissue model and the real-time numerical simulation of its mechanical behavior. Section 4 reviews the anatomical structure of facial muscles, describes the muscle actuators embedded in our facial tissue model, and describes a facial action coding process that controls these muscles to create recognizable expressions. We demonstrate the ability of our model to produce animation at interactive rates. Section 5 concludes the paper.

2 Background

The generation of synthetic 3D facial faces has been the subject of investigation by computer graphic animators [19]. They have used a number of different techniques to create facial expression and speech. The first attempts involved keyframing where two or more complete facial contortions are established and inbetween positions computed by interpolation [16]. The immense pose space of the human face makes this approach extremely cumbersome for 3D facial animation. This prompted Pars and others to develop parametric models for facial animation [17, 2, 1, 18].

Using parameterized models, animators can create facial expressions by specifying appropriate sets of parameter value sequences; for instance, by interpolating the parameters rather than directly keyframing face shapes. The parameters of these models shaped facial features, such as the mouth, by specifying lip opening height, width, and protrusion. But unless the animator specifies the shape parameters with care the model will produce incorrect shapes, unrealistic motions, and other undesirable effects.

This drawback of these earlier models prompted a movement towards models whose parameters are based on the anatomy of the human face [23, 24, 31, 30, 29]. Such models can produce the desired effects using a relatively small number of parameters based on facial muscle structures. When anatomically-based models incorporate facial coding schemes [7] as control procedures, it becomes relatively straightforward to synthesize a range of recognizable expressions.

The purely geometric nature of prior face models ignore many of the complexities of human facial tissue. These models consider the skin as an infinitesimally thin surface with no underlying structure, and deformations are generated by geometrically manipulating the surface [13, 29]. Consequently, attempts to mimic many of the subtleties of facial tissue deformation, such as wrinkles and furrows, are spurious.

This paper suggests that the physically-based approach is fundamentally superior. A wealth of biomechanical literature on tissue mechanics [12] has provided motivation for finite element models of facial tissue suitable for surgical simulation [14, 4]. Dong [4] simulates and analyzes the closure of skin excisions on an idealized three layered model of facial tissue. The computer simulation mimics "dog ears" (a raised area of skin that tends to form at the ends of a closed wound), and the closing of excisions traversing Lauger's lines. Facial muscle articulations, that generate skin deformations, was not the focus of her thesis. However, investigations by Pieper [21] and Waters [32] describe facial
tissue models constructed from deformable lattices, a simple class of discrete deformable models [27], capable of deformations under muscular control.

This paper continues the physically-based approach and presents an applied facial tissue model capable of realistic facial expression under the control of an anatomically-based muscle articulation process.

3 A Physically-Based Tissue Model

This section takes the point of view that facial tissue is more than a lar control. Along Langer's lines, than across them skin is nonhomogeneous and nonisotropic (lower stiffness along Langer's lines than across them) like structure, skin is nonhomogeneous and nonisotropic (lower stiffness along Langer's lines than across them).

3.1 Histology and Mechanics of Facial Tissue

Human skin has a layered structure consisting of the epidermis, dermis, and subcutaneous ground substance. Because of its layered, network-like structure, skin is nonhomogeneous and nonisotropic (lower stiffness along Langer's lines than across them) [12, 14].

The epidermis is a superficial layer of dead cells, which is about one tenth the thickness of the dermis that it protects. The dermis is primarily responsible for the mechanical properties of skin. Dermal tissue is composed of collagen (72%) and elastin (4%) fibers forming a densely convoluted network in a gelatinous ground substance (26%). Under low stress, dermal tissue offers very low resistance to stretch as the collagen fibers begin to uncoil in the direction of the strain, but under greater stress the fully uncoiled collagen fibers resist stretch much more markedly. This yields a biphasic stress-strain curve Figure. 1. The incompressible ground substance retards the motion of the fibers and gives rise to time-dependent viscoelastic behavior: stress relaxation at constant strain, strain creep at constant stress, and hysteresis under cyclic loading (see [28] for a discussion of viscoelastic materials and related graphics models). Finally, the elastin fibers act like elastic springs which return the collagen fibers to their coiled condition under zero load.

Underneath the skin is a layer of subcutaneous fatty tissue which allows the skin to slide rather easily over fibrous fascia covering the underlying muscle layer (see Section 4).

3.2 Deformable Meshes

A deformable lattice is a type of discrete deformable model constructed from point masses connected by springs [9]. Figure 2 shows an example of a 3D deformable lattice constructed from hexahedral elements.

Let node \( i \), where \( i = 1, \ldots, N \), be a point mass \( m_i \) and 3-space position \( x_i(t) = [x_i(t), y_i(t), z_i(t)]' \). The velocity of the node is \( v_i = dx_i/dt \) and its acceleration is \( a_i = d^2x_i/dt^2 \).

Let spring \( k \) have natural length \( L_k \) and stiffness \( c_k \). Let us assume the spring connects node \( i \) to node \( j \), where \( r_k = x_i - x_j \) is the vector separation of the nodes. The actual length of the spring is \( L_k = |r_k| \).

The deformation of the spring is \( \varepsilon_k = L_k - L_k \). Then, the force the spring exerts on node \( i \) is defined as

\[ f_k = c_k \varepsilon_k \]

The spring force is a nonlinear function of node positions because the vector norm \( L_k \) involves roots of sums of squares.

The total force on node \( i \) due to springs which connect it to other nodes \( j \in N_i \) in the deformable lattice is

\[ f_i(t) = \sum_{j \in N_i} c_{ij} \varepsilon_{ij} \]

where the small-strain stiffness \( c_k \) is smaller than than the large-strain stiffness \( \beta_k \) (Figure. 1).

3.3 A Trilayer Model of Facial Tissue

The deformable lattice in Figure. 2 will collapse in the absence of support because although each hexahedral element resists extension and compression forces, it does not resist twist and shear forces. If the faces of each element are cross-strutted with springs, the lattice will resist twisting and shearing, yielding a structurally stable model. An advantage of tetrahedral and pentahedral elements over hexahedra is that they achieve structural stability with fewer springs. We use combinations of structurally stable elements to construct deformable lattice models of facial tissues.

Figure 3 illustrates the geometry of a facial tissue model consisting of three layers of elements representing the cutaneous tissue, subcutaneous tissue, and muscle layer. The springs (line segments) in each layer have different stiffness parameters in accordance with the non-homogeneity of real facial tissue. The topmost surface represents the epidermis (a rather stiff layer of keratin and collagen) and we set the spring stiffnesses so as to make it moderately resistant to deformation. The biphasic springs underneath the epidermis represent the dermis. The springs in the second layer are highly deformable, reflecting the nature of subcutaneous fatty tissue. Nodes on the bottommost surface of the second layer represent the fascia to which the muscle fibers are attached. The bottom surface are fixed in place in bone.

To account for the incompressibility of the cutaneous ground substance and the subcutaneous fatty tissues, we include a constraint into each element which minimizes the deviation of the volume of a deformed element from its natural volume at rest. Differentiation of this constraint yields a volume preservation force \( q_i \) which acts on each node in the lattice from the volumes centered.
To create a more accurate model, we define a biphasic spring which, like dermal tissue, is readily extensible at low strains, but exerts rapidly increasing restoring stresses after reaching a certain strain \( \varepsilon^c \). The biphasic spring exerts a force with a stiffness

\[
e_k = \begin{cases} 
\alpha_k & \text{when } \varepsilon_k \leq \varepsilon^c; \\
\beta_k & \text{when } \varepsilon_k > \varepsilon^c; 
\end{cases}
\]

where the small-strain stiffness \( \alpha_k \) is smaller than than the large-strain stiffness \( \beta_k \) (Figure 1).

### 3.3 A Trilayer Model of Facial Tissue

The deformable lattice in Figure 2 will collapse in the absence of support because although each hexahedral element resists extension and compression forces, it does not resist twist and shear forces. If the faces of each element are cross-strutted with springs, the lattice will resist twisting and shearing, yielding a structurally stable model. An advantage of tetrahedral and hexahedral elements over hexahedra is that they achieve structural stability with fewer springs. We use combinations of structurally stable elements to construct deformable lattice models of facial tissues.

Figure 3 illustrates the geometry of a facial tissue model consisting of three layers of elements representing the cutaneous tissue, subcutaneous tissue, and muscle layer. The springs (line segments) in each layer have different stiffness parameters in accordance with the non-homogeneity of real facial tissue. The topmost surface represents the epidermis (a rather stiff layer of keratin and collagen) and we set the spring stiffness \( a_k \) for each node in the epidermis which represents the dermis.

The springs in the second layer are highly deformable, reflecting the viscoelastic behavior of real skin. There are several ways to include nonlinear viscoelasticity into the elastic spring model (see Figure 3, which is analogous to placing a dashpot in parallel with each spring, so our simulated facial tissue model exhibits a somewhat viscoelastic response.

### 4 Controlling Facial Articulation

To simulate the dynamics of the adaptive mesh, we provide initial positions \( x_i^0 \) and velocities \( v_i^0 \) for each node \( i \) for \( i = 1, \ldots, N \), and numerically integrate the equations of motion forward through time until the mesh stabilizes: \( v_i = a_i = 0 \).

At each subsequent time step, \( \Delta t, 2\Delta t, \ldots, t + \Delta t, \ldots \) we evaluate the current forces, current accelerations, new velocities, and new positions using the explicit Euler time-integration procedure [25]:

\[
\begin{align*}
a_i^{t+\Delta t} &= \frac{a_i^t - \gamma_i v_i^t - q_i^t}{m_i} \\
v_i^{t+\Delta t} &= v_i^t + \Delta t a_i^{t+\Delta t} \\
x_i^{t+\Delta t} &= x_i^t + \Delta t v_i^{t+\Delta t}.
\end{align*}
\]

The stability of the Euler procedure is inherently limited, but its convergence is facilitated by the overdamped dynamics of the mesh and the high flexibility of the biphasic springs in the small-strain region (see [26] for a discussion of the effect that deformable model flexibility has on the numerical procedure). However, modest increases in the size of the time step or too large a value for the large-strain stiffness \( \beta_k \) will cause divergence.

It is possible to maintain convergence by using a more stable numerical procedure such as the Adams-Bashforth-Moulton method [25], but the associated computational complexity per time step will preclude interactive performance. As a compromise solution to the stability/complexity tradeoff, we chose a second-order Runge-Kutta method which requires two evaluations of the forces per time step.

We point out in passing that the spring force equations do not model the viscoelasticity of real skin. There are several ways to include viscous behavior into the elastic spring model (see [28]). Note, however, that the explicit time integration procedure introduces an artificial stiffness which is analogous to placing a dashpot in parallel with each spring, so our simulated facial tissue model exhibits a somewhat viscoelastic response.

### 3.4 Numerical Simulation

The discrete Lagrangian equations of motion for the dynamic node/spring system is the system of coupled, second order ordinary differential equations

\[
\begin{align*}
m_i \ddot{x}_i + \gamma_i \dot{x}_i + q_i &= f_i; \\ &= 1, \ldots, N.
\end{align*}
\]

where \( \gamma_i \) is a velocity-dependent damping coefficient which dissipates kinetic energy in the lattice through friction. For the most part, facial tissue exhibits a slightly overdamped behavior.

### 4 Controlling Facial Articulation

To simulate the dynamics of the adaptive mesh, we provide initial positions \( x_i^0 \) and velocities \( v_i^0 \) for each node \( i \) for \( i = 1, \ldots, N \), and numerically integrate the equations of motion forward through time until the mesh stabilizes: \( v_i = a_i = 0 \).

At each subsequent time step, \( \Delta t, 2\Delta t, \ldots, t + \Delta t, \ldots \) we evaluate the current forces, current accelerations, new velocities, and new positions using the explicit Euler time-integration procedure [25]:

\[
\begin{align*}
a_i^{t+\Delta t} &= \frac{a_i^t - \gamma_i v_i^t - q_i^t}{m_i} \\
v_i^{t+\Delta t} &= v_i^t + \Delta t a_i^{t+\Delta t} \\
x_i^{t+\Delta t} &= x_i^t + \Delta t v_i^{t+\Delta t}.
\end{align*}
\]

The stability of the Euler procedure is inherently limited, but its convergence is facilitated by the overdamped dynamics of the mesh and the high flexibility of the biphasic springs in the small-strain region (see [26] for a discussion of the effect that deformable model flexibility has on the numerical procedure). However, modest increases in the size of the time step or too large a value for the large-strain stiffness \( \beta_k \) will cause divergence.

It is possible to maintain convergence by using a more stable numerical procedure such as the Adams-Bashforth-Moulton method [25], but the associated computational complexity per time step will preclude interactive performance. As a compromise solution to the stability/complexity tradeoff, we chose a second-order Runge-Kutta method which requires two evaluations of the forces per time step.

We point out in passing that the spring force equations do not model the viscoelasticity of real skin. There are several ways to include viscous behavior into the elastic spring model (see [28]). Note, however, that the explicit time integration procedure introduces an artificial stiffness which is analogous to placing a dashpot in parallel with each spring, so our simulated facial tissue model exhibits a somewhat viscoelastic response.

### 4 Controlling Facial Articulation

The confluence of facial muscle contractions deform the face into what we recognize as expressions. Extensive work by the psychologists of non-verbal communication has established a basic categorization of facial expressions generic to the human race [6]. Ekman, Friesen and Ellsworth [6] propose happiness, anger, fear, surprise, disgust/contempt and sadness as the six primary affect categories. Other expressions such as, instead, clam, bitterness, pride, irony, insecurity and skepticism can be displayed on the face but have not been as firmly established.

Ekman and Friesen have also developed the Facial Action Coding System (FACS), a quantified abstraction of the actions of facial muscles, as a means of recording facial expressions independent of cultural or personal interpretation [7]. The FACS represents facial expressions in terms of 66 action units (AU), which involve one or more muscles, and associated activation levels. The AUs are grouped into those that affect the upper face and the lower face, and they include vertical actions, horizontal actions, oblique actions, orbital actions, and miscellaneous actions such as nostril shape, jaw drop, and head and eye position.

The FACS can be used as reference for the development of a parametrized facial muscle process [31, 30] which also provides a valuable means to create grouped functional synergies. For the control of facial articulations in this research we have developed six levels of abstraction based on muscle activation. From highest to lowest, the levels of abstraction encompassed by our model are:

1. \textit{Expressions.} The animator normally interacts with the face model at the expression level. The model will execute commands to synthesize any of the six canonical expressions within a specified time interval and to a given degree of markedness.

2. \textit{Control.} A control process translates expression level instructions to muscle action units. Action units actuate individual muscles and coordinate groups of muscles to perform the expression.

3. \textit{Muscles.} The muscle model is the basic activation mechanism of the face model. Each muscle consists of a bundle of muscle fibers that, by contracting, apply forces that distort the facial tissue.
a menu driven interface which allows the specification of individual muscle contractions and the rotation of the jaw. The epidermis layer is Gouraud shaded and the eyes and teeth positioned and rendered. Figure 6 and Figure 7 illustrates the facial skin model at rest while Figure 8 has the jaw rotated. Figure 9 and 10 demonstrate asymmetric contractions on the syzygomatic major, frontalis and the labii superioris muscles. Figure 11 illustrates an expression of happiness. Figure 12 has the jaw rotated. Figure 13 illustrates the expression of anger. Note the wrinkles in the skin that appear around the frowning mouth, the nasolabial furrows, and the bulging of the inner portion of the brow at the root of the nose. This effect, although difficult to accomplish with a purely geometric model, emerges automatically through the physical simulation of the facial tissue, principally due to the volume preservation constraints.

5 Conclusion

This paper has presented a trilayer model of human facial tissue. The model uses physically-based techniques and a set of anatomically motivated facial muscle actuators. Despite its sophistication, the model is efficient enough to produce facial animation at rapid screen refresh rates on a high performance graphics workstation. The facial model can also generate a wide range of facial expressions which can be correlated with a facial coding system.

Acknowledgements

We thank Reid Smith for supporting our research in facial modeling at SLCS and Jahir Pabon for valuable future research suggestions.

References

4. Physics. This level implements a physically-based model of human facial tissue. The tissue model is a layered deformable lattice. The activated muscle fibers displace nodes from their resting positions, thus propagating stresses which deform the synthetic tissue.

5. Geometry. Standard numerical techniques simulate the physical effects of muscle induced stresses on the tissue. These stresses distort the face away from its neutral geometry to an expressive geometry. The geometric representation of the face is a nonuniform mesh of polyhedral elements.

6. Images. Standard rendering techniques, accelerated by special purpose hardware, transform the deformed facial geometry, along with viewpoint, lightsource, and reflectance information into synthetic images for visualization.

4.1 Facial Muscle Anatomy

Facial muscles can exert traction on the facial tissue to create expressions. When the muscles contract, they usually pull the facial soft tissue to which they attach towards the place where they emerge from the underlying skeletal framework.

Muscles are bundles of fibers working in unison. Shorter fibers are more powerful but have a smaller range of movement than longer fibers. The shape of the fiber bundle determines the muscle type and its functionality. There are three main types of facial muscles, linear, sphincter, and sheet. Linear muscle, such as the zygomaticus major which attaches to and raises the corner of the mouth, consists of a bundle of fibers which share a common emergence point in bone. Sheet muscle, such as the orbicularis oris which circles the mouth and can purse the lips into a pout, consists of fibers that loop around facial orifices and can draw towards a virtual center; an example is the orbicularis oris which circles the mouth and can purse the lips into a pout.

4.2 Synthetic Muscle Effectors

The model incorporates the FACS representation implemented as part of Water's prior geometric model [31]. The FACS abstraction suppresses the low-level details of muscle motions, allowing the facial animator to think in terms of intermediate-level tasks and skills or higher-level expressions. Although prohibitively tedious to simulate all the muscles in the human face, Waters [31] and others have achieved a broad range of facial expression using on the order of 20 muscle actuators. These actuators run through the bottom layer of the trilayer tissue model. Muscles fibers emerge from nodes fixed in bone at the bottom of the layer and attach to mobile nodes on the upper surface of the layer.

The displacement of node $j$ from $x_j$ to $x_j'$ due to muscle contraction is a weighted sum of all the muscle displacements, for $i = 1, ..., N_i$, acting on node $j$: $x_j' = x_j + \sum_{i=1}^{N_i} c_{ij} h_j m_i$, (1)

where $c_{ij}$ are weighted muscle activations, and where $m_i$ is the muscle rest length, $r_{ij}$ is the computed distance to the tail of the muscle vector, and $h_j$ is the muscle blend

$$m_i = m_i^0 - m_i^2$$

$$r_{ij} = m_i^0 - x_j$$

$$h_j = \begin{cases} \cos \left( \frac{\pi r_{ij}}{a_i} \right) & \text{for } |r_{ij}| \leq a_i \\ 0 & \text{otherwise} \end{cases}$$

where $a_i$ is the radius of influence of the cosine blend profile.

Once all the muscle interactions have been computed, the lattice nodes $x_j$ are displaced to their new positions $x_j'$. As a result, those nodes not influenced by the muscle contraction are in an unstable state, and unbalanced forces propagate through the lattice to establish a new equilibrium position. Figure 4 illustrates the stable rest state after the contraction of a muscle.

4.3 Assembling the Face Model

An automatic procedure assembles the physically-based 3D facial model from a nonuniform triangular mesh of a face. The nodes and springs of the initial mesh represent the epidermis. Normal vectors from the center of gravity of each triangle are projected below the surface of the face to establish nodes at the subcutaneous tissue interface. Tetrahedral units are then constructed by attaching springs from the epidermis triangular nodes to the new nodes. The new springs represent the dermal layer. Short weak springs are then attached from the second layer to another set of nodes to form the subcutaneous layer. A final set of springs are attached from these nodes and anchored at the other end to the bone, thus establishing the muscles. The muscle fibers are then automatically inserted through the muscle layer from their emergence in bone to their nodes of attachment which are computed automatically using the equations given in the prior section.

Figure 4 Flesh lattice with embedded muscle. (a) Initial geometry with muscle relaxed. (b) Equilibrium position with muscle contracted and lower layers suppressed for clarity. (1,802 springs)

Figure 5 3D face model. (a) Undeformed geometry of the skin layer. (b) Deformed geometry under the influence of AU1 and AU12 (only epidermis is displayed for clarity).

We have constructed a physically-based face model starting from the topology and geometric information of Water's prior geometric model [31]. Figure 5 shows the skin topology after lattice construction. A total of 960 polygons are used in the skin model, which results in approximately 6,500 spring units. Figure illustrates the skin distortion after the influence of the zygomaticus major muscle (AU12), which raises the corner of the lip, and the inner frontalis muscle (AU1) which raises the inner portion of the brow.

4.4 Examples

The physically-based facial model is efficient enough to simulate and render at screen refresh rates higher than 8Hz on a single chip of a Silicon Graphics 4D/240GTX workstation. Figures 6-14 illustrate a variety of stills from an interactive session. The user modified the model through