

## RESEARCH ARTICLE

# Physically-based forehead animation including wrinkles

Mark Warburton\* and Steve Maddock

Department of Computer Science, The University of Sheffield, 211 Portobello, Sheffield, S1 4DP, UK

## ABSTRACT

Physically-based animation techniques enable more realistic and accurate animation to be created. We present a fully physically-based approach for efficiently producing realistic-looking animations of facial movement, including animation of expressive wrinkles. This involves simulation of detailed voxel-based models using a graphics processing unit-based total Lagrangian explicit dynamic finite element solver with an anatomical muscle contraction model, and advanced boundary conditions that can model the sliding of soft tissue over the skull. The flexibility of our approach enables detailed animations of gross and fine-scale soft-tissue movement to be easily produced with different muscle structures and material parameters, for example, to animate different aged skins. Although we focus on the forehead, our approach can be used to animate any multi-layered soft body. © 2014 The Authors. *Computer Animation and Virtual Worlds* published by John Wiley & Sons, Ltd.

## KEYWORDS

physically-based animation; facial animation; soft-tissue animation; wrinkle animation; finite element method

### \*Correspondence

Mark Warburton, Department of Computer Science, The University of Sheffield, 211 Portobello, Sheffield, S1 4DP, UK.

E-mail: M.Warburton@dcs.shef.ac.uk

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

[Copyright line has been changed since first online publication 26 November 2013].

## 1. INTRODUCTION

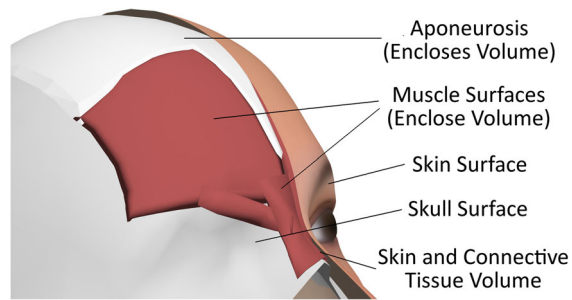
Facial modelling and animation is one of the most challenging areas of computer graphics. Currently, most facial animation requires performance-capture data, or models to be manipulated by artists. However, by using a physically-based approach, the effects of muscle contractions can be propagated through the facial soft tissue to automatically deform the model in a more realistic and anatomical manner.

Physics-based soft-tissue simulation approaches often focus on either efficiently producing realistic-looking animations for computer graphics applications [1,2] or simulating models with high physical accuracy for studying soft-tissue behaviour [3,4] or surgical simulation [5,6]. The former simulates large areas, such as the face, with just enough physics to efficiently produce the desired realism of animations. The skin and connective tissue are normally simulated using the efficient mass-spring (MS) method [2,7], or a physics engine that focusses on performance and stability [8,9]. Simpler and more efficient heuristic techniques are often used to animate muscles represented as vectors [10,11] or geometric volumes [2,12] and layer

skin wrinkles onto a surface mesh by modifying the mesh texture [13,14] or geometry [14,15]. Such techniques have also been used to layer wrinkles onto MS facial models [16,17].

On the other hand, applications with high accuracy requirements often use much more detailed physics-based models accurately to simulate small areas, such as a block of skin for wrinkle simulation [4,18], or they involve small deformations [5,19]. These approaches are normally required to use the more accurate but computationally complex finite element (FE) method [3,6], or the FE-based but precomputation-heavy mass-tensor (MT) method [20,21]. When modelled, they often simulate muscles [18,22] and wrinkles [3,4] in a physics-based manner. Given increases in computational power, and the use of graphics processing unit (GPU) computing architectures, complex FE simulations are now possible in real time [23].

By modelling a more physics-based behaviour than current computer graphics approaches, the aim of this research is to develop a fully physically-based approach for efficiently producing realistic-looking animations of facial movement, including animation of expressive wrinkles, focussing on the forehead region of the face. This involves



**Figure 1.** Surfaces and volumes of a forehead soft-tissue model. The volume between the skull and skin surfaces is discretised and simulated.

simulation of forehead soft-tissue models (Figure 1) using the accurate non-linear total Lagrangian explicit dynamic (TLED) FE method with an anatomical muscle contraction model. Advanced boundary conditions constrain nodes and can model the sliding effect between superficial soft-tissue layers, and the deep layers and skull [24], which is often neglected [19,25]. Our approach can also be used to animate any multi-layered soft body (not just soft tissue).

Physics-based simulations require an appropriate simulation model. Either such models can conform to a surface mesh [19], or a non-conforming model [25], such as a voxel representation, can be used [26]. High-quality conforming models are often difficult and time-consuming to create, whereas non-conforming models can enable more efficient production of stable, realistic-looking animations for computer graphics applications. The GPU-based implementation of our animation approach has been optimised to exploit the computational advantages offered with the voxel-based models we use [26,27].

The following sections discuss related work, followed by an overview of our animation approach and an introduction to our model creation process. Details of the simulation process are then given, including the TLED FE solver, muscle contraction model, and processing of boundary conditions, followed by example animations. Finally, a discussion on the current difficulties and issues with our approach and examples is presented, along with possible improvements.

## 2. RELATED WORK

### 2.1. Physically-Based Facial Animation

Physically-based facial animation systems for computer graphics applications normally consist of muscle and skin models, sometimes along with a skull model and wrinkle models. Skull models range from low-resolution offsets from facial surfaces [28] to accurate models from medical data [25], facilitating more anatomical muscle attachments and collision detection. For increased realism of jaw movement, a rotatable mandible can be used [11].

Muscles can be represented as vectors [10,11], although more realistic approaches consider geometric [2,12] or physics-based volumes [29,30]. Many muscle contraction models are based on a Hill-type model [30,31], some of which are biologically inspired [18]. The direction of contraction can be approximated as parallel to a central action curve [32] or, more anatomically, by using a fibre field [25].

Various facial soft-tissue models have been proposed, including physics-engine-based models that focus mainly on efficiency and stability [8,9]. MS models can also be used, and many MS facial models are based on Terzopoulos and Waters' layered model [1], which was modified by Lee *et al.* [28], with various enhancements and different muscle models [7,11]. With MS systems, additional functions to simulate, for example, incompressibility and volume preservation, are necessary as these models by default contain no volume representation [2,12]. Although uncommon owing to their lack of accuracy, the MS method, and variations of it, can be used for surgical applications [33].

More anatomical and realistic facial models have been developed using the FE method [22,25]. Muscle activations can be estimated from performance-capture data to drive FE facial models [25,34]. However, because of its computational cost, the FE method is mainly used with scientific [19,35] or surgical applications [5,6] that require high accuracy, which normally use simple linear isotropic constitutive models to simulate small displacements. The linear MT method has also been used for simulating small displacements with such applications [20,36].

### 2.2. Skin Wrinkle Animation

For computer graphics applications, skin wrinkles are usually layered onto a surface mesh using texture-mapping techniques [13,37], which are computationally efficient and suitable for fine wrinkles, or by using a heuristic technique to modify the geometry of the mesh [15,38], which can produce more realistic-looking wrinkles. Geometric approaches require a high-resolution mesh, although the resolution of meshes can be adaptively refined when required during animations [39]. A combination of texture-mapping techniques for fine wrinkles and geometric techniques for large wrinkles can also be used [14,16]. Some approaches require the wrinkles to be specified by the user [40], which can be blended between poses, whereas others are calculated automatically [41], although with less anatomical accuracy. Performance-capture techniques can also be used to produce wrinkle animations [42,43], which require the facial or wrinkle model to be configured to the performer.

Wrinkling approaches have also been used with physics-based facial models; for example, Zhang *et al.* developed a geometric technique to layer wrinkles on an MS facial model based on the gross deformation rather than simulate them in a physics-based manner [17]. Wu *et al.* modelled large and fine static and dynamic wrinkles on

a simple biomechanical facial model using both texture-mapping and geometric techniques [16], and plasticity was also considered to simulate ageing.

### 2.3. Detailed Soft-Tissue Simulation

Multi-layered FE models of small areas of soft tissue have been developed for accurate simulation of expressive or ageing wrinkles [4,44], some of which simulate factors such as anisotropy and viscoelasticity [3]. Complex constitutive models have also been proposed specifically to simulate soft tissue [45], which are able to simulate complex skin properties such as hysteresis and preconditioning. Because of its efficiency, the TLED algorithm has been used for various non-linear FE soft-tissue simulations of biological organs [46], including GPU implementations [23], resulting in large speed-ups. The MT method [21] and the less accurate MS method [47] and physics engines [48] have also been used for such soft-tissue simulations.

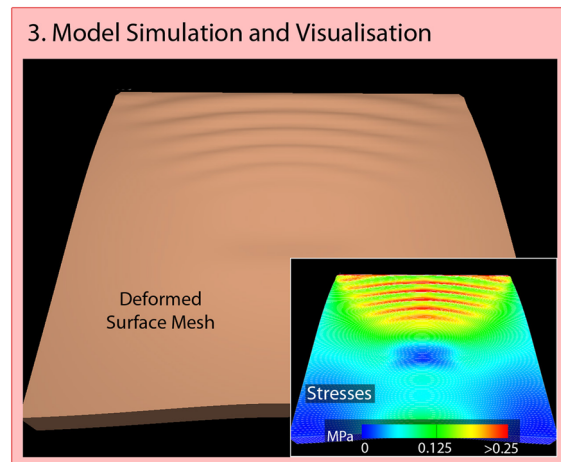
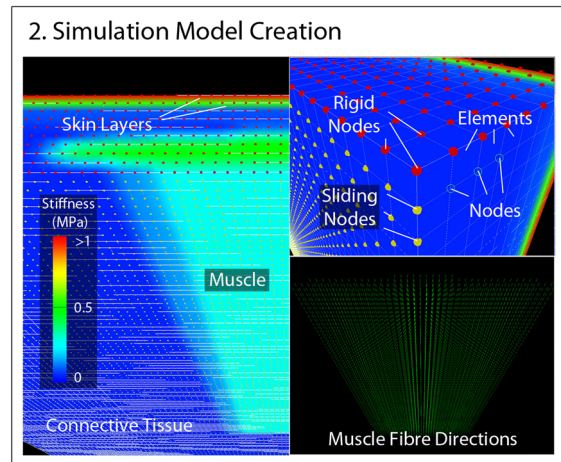
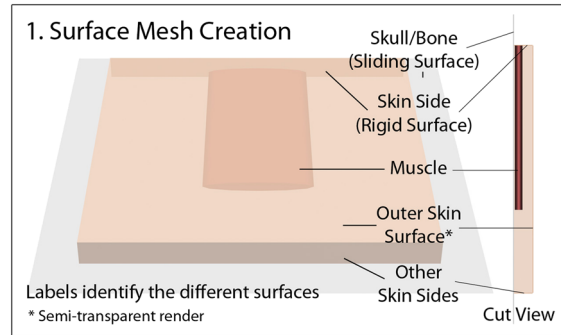
Rather than modelling a small area, we focus on using detailed FE facial models. However, unlike current FE facial animation approaches [25,35], we also model detail such as skin layers using volumetric FEs, which are necessary to simulate fine details such as wrinkles [3], in order to produce both realistic-looking gross and fine-scale facial movements in a fully physics-based manner. Our simulation models are also optimised for GPU simulation. In addition to computer graphics applications, owing to the detail and accuracy of the simulations, our animation approach could also be useful in other fields, such as biomechanics and surgery.

## 3. ANIMATION PROCESS OVERVIEW

Figure 2 shows an overview of our entire animation process, which involves three major stages:

- (1) Creating the surface mesh for an object
- (2) Creating a suitable simulation model
- (3) Simulating and visualising the model over time

The surface mesh can be created using any three-dimensional modelling software, and it can contain various surfaces and volumes (Figure 1). Our model creation process is then used to automatically discretise the volumes enclosed by this mesh into a collection of nodes that are connected to form elements and compute FE model parameters to produce a simulation model [27]. We use non-conforming (voxel-based) hexahedral models because of model creation, performance, and stability advantages [26], and surface meshes are bound to these for visual purposes. The final stage of the animation process (the focus of this paper) involves simulating and visualising the models using our GPU-based FE simulation and visualisation system. The following sections introduce our model creation process and present the details of our simulation process.



**Figure 2.** The stages of our animation process. Model simulation is the focus of this paper.

## 4. MODEL CREATION

The full details of our model creation process have been previously presented [27]. This process automatically creates animatable voxel-based FE simulation models of facial

soft tissue, including computation of all model properties. This involves five main stages:

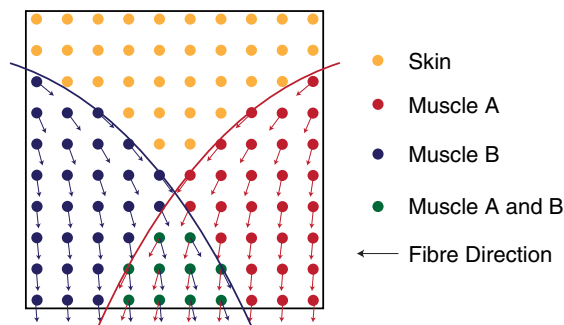
- (1) Voxelising the volumes of the multi-surface mesh
- (2) Computing skin layers and element material properties
- (3) Computing element muscle properties
- (4) Computing boundary conditions, such as rigid or constrained nodes
- (5) Binding the surface mesh to the simulation model

As shown by Figure 3, a sampling procedure is used to approximate the proportions of overlap between voxels (used as hexahedral elements) and volumes of the surface mesh. These proportions are used to compute element material properties and weight element muscle stresses. Constant thickness layers can be created from the outer skin surface. Fibre directions are also computed at samples using the gradient of non-uniform rational B-spline volume muscle approximations with respect to the length of the muscle [49], and these are used to compute element muscle fibre directions.

To define the model boundary conditions, it is possible to restrict the movement of particular nodes, as described in Section 5.4. Such nodes are identified given a collection of internal and external boundary surfaces; for example, nodes that slide over the external skull surface are identified as the outermost nodes that approximate this surface. The final stage of the model creation process binds the vertices of the surface mesh to the elements of the simulation model, enabling them to be animated using trilinear interpolation and extrapolation. Although the focus of this paper involves soft-tissue models, the process can be used to create any multi-layered model from any closed surface meshes.

## 5. MODEL SIMULATION

Our Compute Unified Device Architecture (CUDA)-based FE simulation system uses the non-linear TLED formulation of the FE method [46], which has various



**Figure 3.** Voxel element samples assigned a particular material and a fibre direction for each overlapping muscle. Note that this is a two-dimensional illustration of a three-dimensional process.

advantages for simulating soft tissue. Because of the requirement of a small but efficient simulation timestep, explicit time integration is highly suited for simulating the complex non-linear material behaviour of soft tissue under large deformations, as well as contact and sliding. Explicit methods are also inherently parallel, making them suitable for GPU implementation. To further increase simulation efficiency, the total Lagrangian (TL) formulation enables some variables to be precomputed. Being dynamic, inertial and damping effects are also considered.

The procedure to compute a timestep using our simulation system is summarised by Algorithm 1. The iterations within each loop are independent and are therefore computed in parallel on the GPU. The hexahedral elements of our simulation models are treated as reduced-integration eight-node hexahedra, which have a single central integration point at which element strains and stresses are evaluated. Factors such as thermal effects and fluid dynamics are not considered necessary for the level of accuracy we desire. The following subsections present details of our simulation approach.

**Algorithm 1:** The process to compute a timestep for non-conforming elements.

```

1 foreach element, m do
2   foreach integration point do
3     foreach muscle overlapping m do
4       | Calculate active muscle stresses;
5     foreach material inside m do
6       | Calculate material stresses;
7     | Calculate internal nodal forces;
8   | Calculate hourglass forces;
9 foreach node, n do
10  | if n not rigid then
11  | | Calculate nodal displacements;
12 foreach sliding node do
13 | | Update nodal displacements;
    
```

### 5.1. The Total Lagrangian Explicit Dynamic Finite Element Method

The full formulation of the dynamic TL FE formulation with explicit time integration has been presented in a previous work [46]. As the TL formulation calculates deformation with respect to the initial object configuration, the Green–Lagrange strain tensor,  ${}^t_0\mathbf{E}$  (at time  $t$  with respect to the initial configuration), is used, with the work-conjugate second Piola–Kirchhoff stress tensor,  ${}^t_0\mathbf{S}$ , to handle the large deformations and rotations that can occur. By using these, the principle of virtual work, the basis of the displacement-based FE method, can be defined as

$$\int_{0V} {}^t_0\mathbf{S}'_0 \delta \mathbf{E} dV = \int_{0V} {}^t \delta \mathbf{u}' \mathbf{f}^{(B)} d^0V + \int_{0S} {}^t \delta \mathbf{u}' \mathbf{f}^{(S)} d^0S \quad (1)$$

where  ${}^t\mathbf{u}$  are displacements,  ${}^t\mathbf{f}^{(B)}$  and  ${}^t\mathbf{f}^{(S)}$  are external body and surface forces, respectively, and  ${}^0V$  and  ${}^0S$  are the volume and surface area over which the forces are applied, respectively.

By discretising an object into elements and representing the variation of displacement over elements using shape functions, strain and stress vectors can be calculated. By relating element strains to nodal displacements, the strain–displacement matrix at an element integration point,  ${}^t_0\mathbf{B}_L$ , can be defined as

$${}^t_0\mathbf{B}_L = ( {}^t_0\mathbf{B}_{L_1} \quad {}^t_0\mathbf{B}_{L_2} \quad \cdots \quad {}^t_0\mathbf{B}_{L_n} ) \quad (2)$$

$${}^t_0\mathbf{B}_{L_a} = {}^0\mathbf{B}_{L_0a} {}^t_0\mathbf{X}^T \quad (3)$$

where  ${}^0\mathbf{B}_{L_0}$  is the initial strain–displacement matrix, constructed from the shape function derivatives,  $n$  is the number of element nodes, and  ${}^t_0\mathbf{X}$  is the deformation gradient. Evaluation of stresses at integration points is performed using the strains and relevant constitutive equation. As we use non-conforming models, the stress vector for an integration point is a weighted sum of the stress vectors calculated for each material used by the element.

By using the shape functions and strain and stress vectors and applying equation 1 for each degree of freedom, this principle of virtual work becomes the following equation of motion:

$$\mathbf{M}^t\ddot{\mathbf{u}} + \mathbf{C}^t\dot{\mathbf{u}} + k({}^t\mathbf{u}) {}^t\mathbf{u} = {}^t\mathbf{r} \quad (4)$$

where  $\mathbf{M}$  is the mass matrix,  $\mathbf{C}$  is the damping matrix,  $k({}^t\mathbf{u})$  is the stiffness matrix, and  ${}^t\mathbf{r}$  is the vector of external forces. Here, forces due to inertia and damping are separated from the external body forces. We use a lumped-mass approximation with mass-proportional Rayleigh damping. Nodal force contributions for each element,  ${}^t\mathbf{f}$ , are independently calculated as

$${}^t\mathbf{f} = k({}^t\mathbf{u}) {}^t\mathbf{u} = \int_{{}^0V} {}^t_0\mathbf{B}_L^T {}^t\hat{\mathbf{s}} d {}^0V \quad (5)$$

where  ${}^t_0\hat{\mathbf{s}}$  is the second Piola–Kirchhoff stress vector. To advance simulations, the central difference time-integration method is used.

## 5.2. Hourglass Control

Whereas using reduced-integration elements overcomes the volume-locking limitations that can occur when simulating incompressible material such as soft tissue [23], an hourglass control technique is required to reduce the hourglass effects (zero-energy modes, whereby element deformations occur such that no strains, and hence no forces to resist the deformation, are produced) that can occur, particularly with large strains. We use a stiffness-based method, which adds artificial stiffness to elements to constrain the different hourglass modes based on element

stiffness [50]. An element hourglass force matrix,  ${}^t\mathbf{F}_H$ , is calculated as

$${}^t\mathbf{F}_H = {}^0\mathbf{H}^t\mathbf{U} \quad (6)$$

$${}^0\mathbf{H} = \kappa \cdot k_{\max} {}^0\gamma {}^0\gamma^T \quad (7)$$

where  ${}^0\mathbf{H}$  is the hourglass matrix,  ${}^t\mathbf{U}$  is a matrix of nodal displacements of the element,  $\kappa$  is a user-defined stiffness parameter,  $k_{\max}$  is the maximum element stiffness, and  ${}^0\gamma$  is the matrix of element hourglass shape vectors. Each row of the hourglass force matrix is an element hourglass force vector corresponding to an element node, and these vectors are added to the relevant nodal internal forces.

## 5.3. Muscle Contraction

To enable active muscle stresses to be generated during simulations, a muscle contraction model has been implemented. As muscles are transversely isotropic with preferred deformation in the fibre direction, as well as an active stress component, this direction is used to compute an additional passive stress component.

For a particular muscle, the additional stress produced by the muscle at an integration point,  ${}^t_0\mathbf{S}_M$ , is computed by adding the active stress,  ${}^t_0\mathbf{S}_{M_{act}}$ , and additional passive stress,  ${}^t_0\mathbf{S}_{M_{pas}}$ :

$${}^t_0\mathbf{S}_M = {}^t_0\mathbf{S}_{M_{act}} + {}^t_0\mathbf{S}_{M_{pas}} \quad (8)$$

$${}^t_0\mathbf{S}_{M_{act}} = {}^tJ {}^t\alpha \sigma_{act} f_{act}({}^t\lambda) {}^0\mathbf{d} {}^0\mathbf{d}^T \quad (9)$$

$${}^t_0\mathbf{S}_{M_{pas}} = {}^tJ \sigma_{pas} f_{pas}({}^t\lambda) {}^0\mathbf{d} {}^0\mathbf{d}^T \quad (10)$$

where  ${}^tJ$  is the Jacobian of the deformation gradient;  $\sigma_{act}$  and  $\sigma_{pas}$  are the active and passive muscle stress references, respectively, both weighted by the element overlap with the muscle;  ${}^t\lambda$  is the muscle fibre stretch ratio; and  ${}^0\mathbf{d}$  is the element muscle fibre direction.  ${}^t\alpha \in [0, 1]$  is the muscle

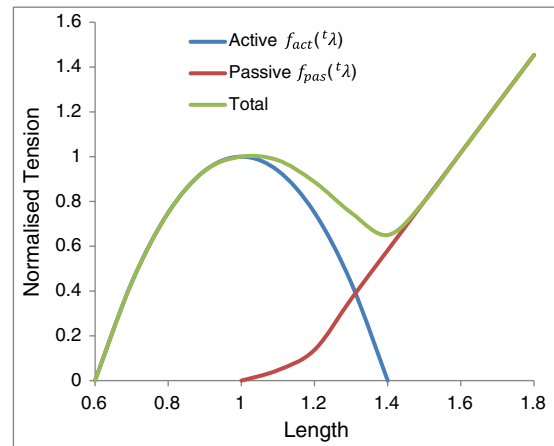


Figure 4. The normalised active,  $f_{act}({}^t\lambda)$ , passive,  $f_{pas}({}^t\lambda)$ , and total muscle tension–length functions.

contraction parameter, which is varied over time according to a function (e.g. linear or cosine) to simulate muscle contraction.  $f_{act}(t\lambda)$  and  $f_{pas}(t\lambda)$  are based on those used by Rörlé *et al.* [30], which, as shown by Figure 4, follow the tension-length properties observed by real muscle. These functions are defined as

$$f_{act}(\lambda) = \begin{cases} -\frac{25}{4\lambda_{opt}^2}\lambda^2 + \frac{25}{2\lambda_{opt}}\lambda - 5.25 & 0.6\lambda_{opt} \leq \lambda \leq 1.4\lambda_{opt} \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

$$f_{pas}(\lambda) = \begin{cases} 0 & \lambda \leq 1 \\ 0.05(\exp^{6.6(\lambda-1)} - 1) & 1 < \lambda \leq \lambda_{pas} \\ 2.18\lambda - 2.47 & \text{otherwise} \end{cases} \quad (12)$$

where  $\lambda_{opt}$  is the optimal fibre stretch, which is normally similar to the stretch at which the passive stress increases sharply,  $\lambda_{pas}$ .

### 5.4. Boundary Conditions

To constrain models during simulations, boundary conditions must be set. With our system, it is possible to set nodes as rigid or sliding (bound by a surface). Rigid nodes are simply fixed with zero displacement throughout simulations and can therefore be used to model muscle attachments on the skull or the roots of retaining ligaments. Sliding nodes are used to model material that is normally attached to but can slide over a rigid surface; for example, in reality, superficial facial soft-tissue layers are able to slide over the stiff deep layers and skull [24], although retaining ligaments normally restrict the separation of these layers. This sliding phenomenon has often been neglected in previous work, in which nodes on the skull are simply treated as rigid nodes [19,25].

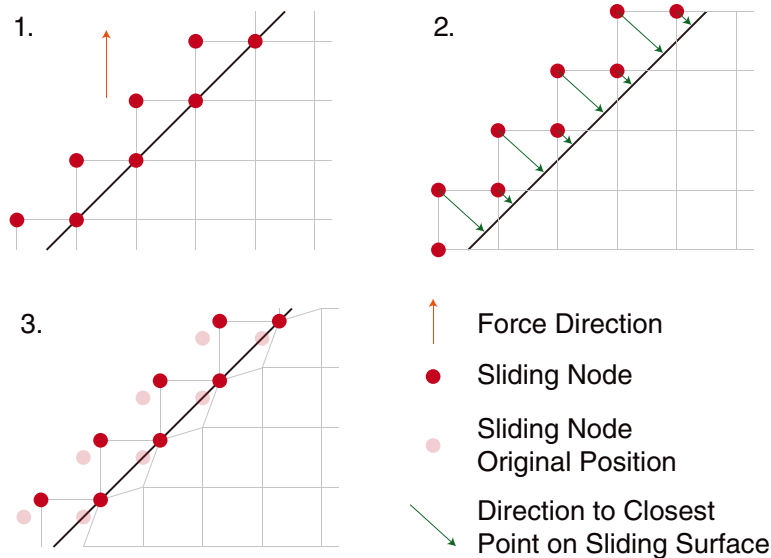
As we are using non-conforming models, sliding nodes would not necessarily lie on the surface they are bound by. To constrain such nodes to follow the shape of the restricting surface, they are forced to maintain a fixed distance away from this surface. Friction is not considered. As shown by Figure 5, the displacement of a sliding node is updated after computation of a timestep to move this node to a position that is distance  $d$  (the fixed distance) away from the sliding surface. This additional displacement,  $\mathbf{u}_S$ , is calculated as

$$\mathbf{u}_S = -b \cdot (d + b\|(\mathbf{p} - \mathbf{c})\|) \cdot \frac{(\mathbf{p} - \mathbf{c})}{\|(\mathbf{p} - \mathbf{c})\|} \quad (13)$$

$$b = \begin{cases} 1 & (\mathbf{p} - \mathbf{c}) \cdot \mathbf{n} < 0 \\ -1 & \text{otherwise} \end{cases} \quad (14)$$

where  $\mathbf{p}$  is the node position,  $\mathbf{c}$  is the closest point on the surface to node, and  $\mathbf{n}$  is the surface normal at this point. To increase computational performance, a GPU-based semi-brute-force broad-phase collision detection algorithm with spatial subdivision has been implemented [51] to prune the number of polygons to be tested when finding the closest surface position for each sliding node.

Because of the fixed-distance restriction with non-conforming elements, our sliding procedure can limit the ability of elements to adapt to the shape of the surface as they move, particularly with lower-resolution models and around highly curved areas. However, these constraints help to suitably restrict our models such that a higher timestep can be used. Alternatively, the penalty method has been used in related work [35], although the produced oscillatory movement can greatly decrease stability, and only penetration is constrained; for example, the movement of soft tissue away from the skull of a facial model would be unrestricted.



**Figure 5.** Sliding nodes moving along a sliding surface. The displacement of sliding nodes before (image 2) and after (image 3) the additional sliding displacement is shown. Note that this is a two-dimensional illustration of a three-dimensional process.

## 5.5. Graphics Processing Unit Implementation

Comprehensive details regarding the GPU implementation of our simulation system have been previously presented [52]. Various GPU implementations of the TLED FE method have been reported [23]; however, our optimised GPU simulation system enables simulations to be displayed as they are computed and is therefore suitable for interactive computer graphics applications, unlike most solvers. Explicit dynamic methods with a lumped-mass approximation and mass-proportional damping can be easily parallelised because of the largely uncoupled equations. For example, the muscle stress (equation 8) and internal nodal force calculations (equation 5) can be carried out independently and in parallel for each integration point, the anti-hourglass calculations (equation 6) for each element, and the nodal displacement (equation 4) and boundary condition calculations (equation 13) for each relevant node.

During a CPU precomputation stage, all simulation values that relate only to the initial configuration, including shape function derivatives, element volumes, unscaled hourglass matrices, and constants in the nodal displacement calculations, are precomputed. All simulation values are then copied into global GPU memory as structures of arrays where possible for efficient memory accesses. By the use of CUDA interoperability with OpenGL, these values and graphics data remain on the GPU throughout the simulation to reduce the amount of slow CPU–GPU data transfer.

As non-conforming elements are used, elements may cross the boundaries of several materials, which may be modelled using the same material model with different parameters or different material models. For each material

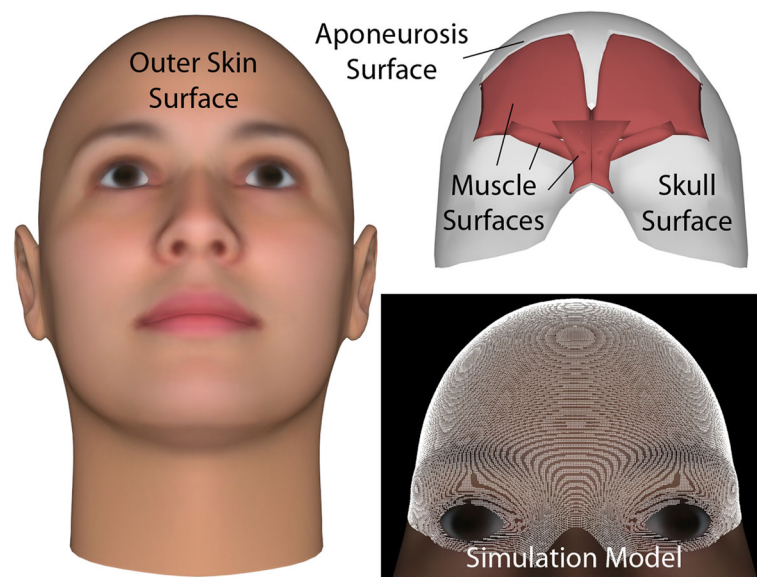
model used by an element, the weight and average parameters are computed in the precomputation stage. These are stored along with a pointer to the relevant GPU function that calculates material passive stresses, which enables easy integration of material types without modifying the kernel that calculates internal nodal forces. The passive stress vector is therefore a weighted combination of the stress vectors calculated for each material model.

For increased computational efficiency, elements are organised by the material models they use to ensure each active thread within a CUDA warp performs the same amount of computation when computing passive stresses. Further optimisations have also been made by computing rows of the internal nodal force vector independently, rather than requiring the entire strain–displacement matrix for an integration point to be stored, and computing the deformation gradients in stages such that element and nodal values are only accessed once.

The simulation system has also been optimised to exploit the computational advantages of using the voxel-based models produced by our model creation process on the GPU [26]. These include both performance and memory advantages resulting from efficient element and node data organisation, facilitating memory coalescing and global memory cache hits, and the reduced amount of element data that is required to be stored, as all elements are initially the same size and shape. This has led to performance increases of almost two times compared with using a conforming hexahedral simulation model.

## 6. RESULTS

Figures 6 and 7 show a forehead model, including the frontalis, procerus, and corrugator supercilli muscles, that



**Figure 6.** Surfaces and the simulation model for a forehead including the frontalis, procerus, and corrugator supercilli muscles.

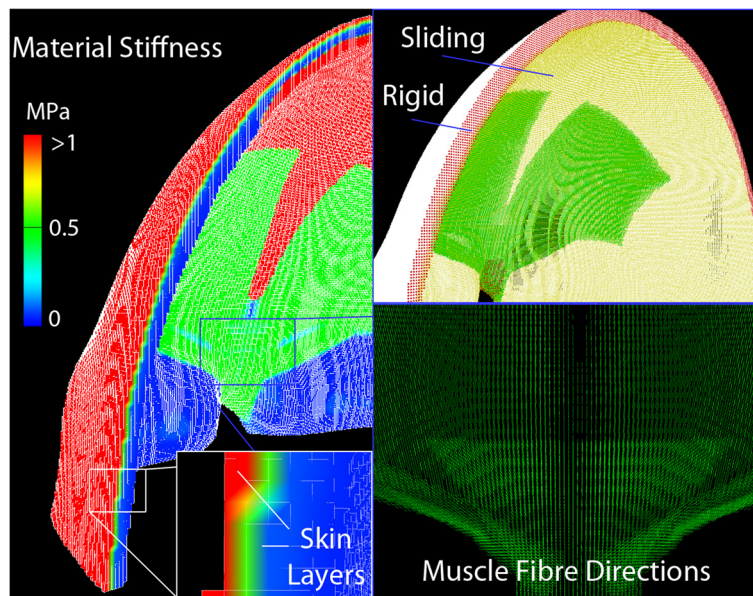


Figure 7. Rear views of the forehead simulation model.

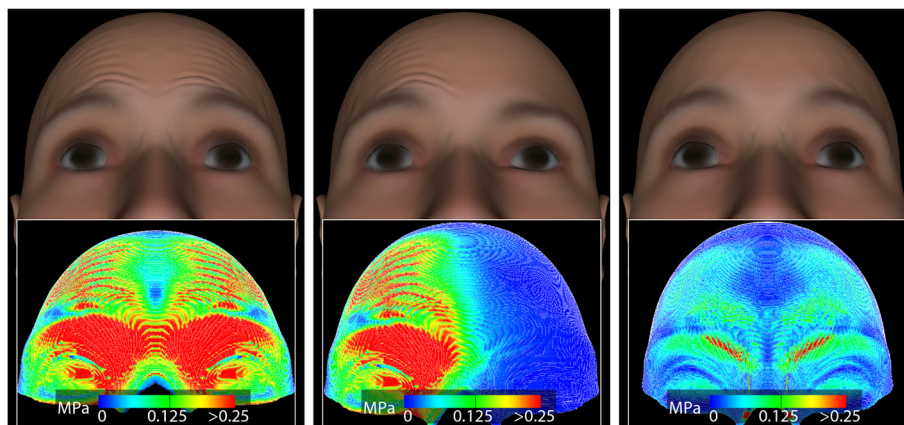


Figure 8. A forehead model under contractions of the frontalis. The example on the right models skin as a single layer (material). Insets show the stresses.

was generated using our model creation process [26]. Figures 8 and 9 show animations using this model, whereas Figures 2 and 10 show animations of flat and curved soft-tissue block models, respectively. Tables I and II show the material properties that were used (based on those reported in literature [4]) and some model statistics. Muscle parameters were estimated from the literature [30] and from testing. Each muscle was assigned an active and passive stress reference of 5 MPa, an optimal fibre stretch of 1 (rest length), and a  $\lambda_{pas}$  (equation 12) of 1.2. As the deep layers are tough and fairly rigid, these were not modelled, and the superficial layers simply slide over the skull or bone surface.

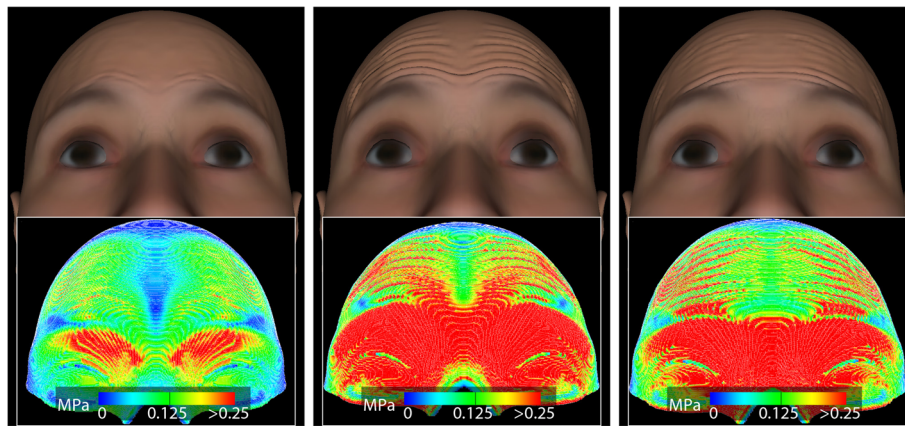
For each example, animations with linear variations of muscle contraction parameters between 0 and 0.75

over 500 ms were produced by varying these parameters over 50 ms and playing back the animations slower using time scaling (Section 7.1). Considering both mass and time scaling, the animations were played back 10 times slower. An estimated mass-proportional damping scale factor of 2 N s/m was used. External forces that have little visual effect on the animations, such as gravity, were neglected.

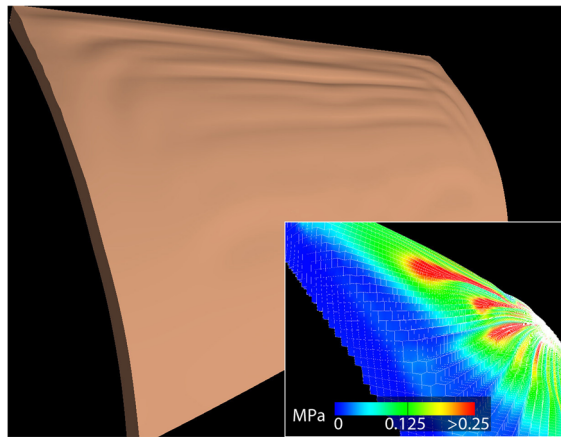
The outer skin surface of the facial mesh was produced using FaceGen,<sup>†</sup> whereas all other surfaces were manually created on the basis of anatomy using three-dimensional modelling tools. The simulation models, both of which

<sup>†</sup><http://www.facegen.com/>





**Figure 9.** Variations of the forehead model under contractions of the frontalis using a lower epidermal stiffness of 24 MPa, representing younger skin (left), using higher active and passive muscle stress references of 50 MPa (middle), and where only the frontalis, and not the procerus and corrugator supercilli muscles, is modelled (right).



**Figure 10.** A soft-tissue-block model with forehead-like muscle structure under contraction of the frontalis.

**Table I.** The neo-Hookean material properties used for the animation examples.

Layer	$\rho^*$ (kg/m <sup>3</sup> )	$E$ (MPa)	$\nu$	Depth (mm)
SC	11 000	48.0000	0.49	0.02
D	11 000	0.0814	0.49	1.80
H	11 000	0.0340	0.49	Remains
M	11 000	0.5000	0.49	~1.00
T	11 000	24.0000	0.49	~1.00

Key:  $\rho$ , density;  $E$ , Young's modulus;  $\nu$ , Poisson ratio; SC, stratum corneum; D, dermis; H, hypodermis; M, muscle; T, tendon.

\*Includes mass scaling.

have constant soft-tissue thickness, were then generated using our model creation system. As the frontalis has no skull attachment, the galea aponeurotica has been modelled on the forehead model (Figure 1) to anchor this muscle and

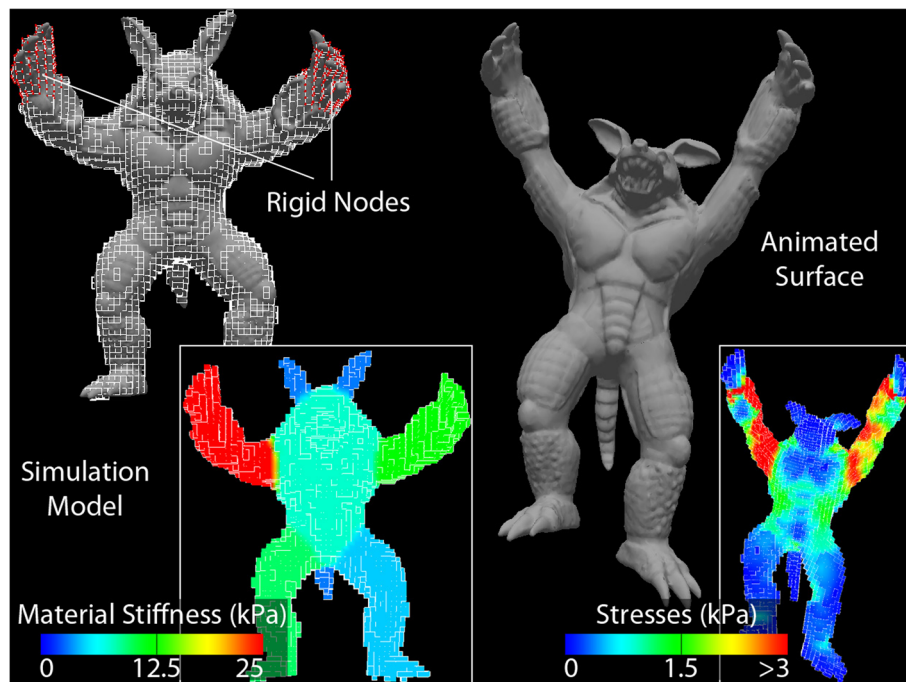
restrict soft-tissue movement towards the top of the head when it contracts. The region of overlap between these structures represents the smooth blend of fibres. Attachments of other muscles were defined by rigid nodes at anatomical attachment locations on the skull.

The animation examples demonstrate the ability of our approach to produce animation of realistic large-scale and fine-scale soft-tissue behaviours, including wrinkles, on the face. When considering skin as a single layer (material), like with many current approaches [19,25], no wrinkles are produced. As shown by Figure 9, by simply changing the material parameters of a model, it is possible to achieve different effects, such as the animation of different aged skins. Also, by using a different muscle structure or a completely different model, different objects can be animated. The animations show the flexibility of our approach to animate different human or non-human soft-tissue structures, producing animations of both large and fine details using an accurate physics-based technique.

Finally, Figure 11 shows an animation of a more generic multi-material object, demonstrating the flexibility of our approach to animate arbitrary multi-material soft bodies. Fewer larger elements were used to animate the gross object movement. Regarding performance, because of the required model complexity to capture the necessary detail for wrinkle simulation, such as skin layers, the soft-tissue animations are not in real time. One bottleneck is the processing of sliding constraints, which contributed to over 40% of the timestep computation time with the forehead simulations, containing 118 276 sliding nodes. Our implementation of this procedure can create an uneven workload between GPU threads and could be further optimised. However, real-time frame rates can be achieved using lower-resolution models for simulating gross deformation, such as the multi-material Armadillo.

**Table II.** Statistics of the examples, using an NVIDIA GTX 680 graphics processing unit (Nvidia Corporation, Santa Clara, CA).

Detail	Face	Flat skin block	Curved skin block	Armadillo
Nodes	629 178	146 410	485 400	19 698
Elements	503 530	129 600	414 868	15 107
Element size (mm <sup>3</sup> )	0.5 <sup>3</sup>	0.5 <sup>3</sup>	0.5 <sup>3</sup>	2.5 <sup>3</sup>
Timestep (ms)	0.005	0.005	0.005	0.15
Timestep computation time (ms)	6.83	1.48	4.04	0.14

**Figure 11.** Animation of a multi-material Stanford Armadillo under gravity. A density of 1250 kg/m<sup>3</sup> was used for the left arm and leg, 750 kg/m<sup>3</sup> for the right arm and leg, and 1000 kg/m<sup>3</sup> elsewhere, while a Poisson ratio of 0.4 was used for each material.

## 7. DISCUSSION

### 7.1. Mass and Time Scaling

With the soft-tissue animations, mass scaling was used to greatly increase the critical timestep, which is roughly proportional to the material density [23]. The actual density of each material ( $\sim 1100 \text{ kg/m}^3$  [19]) was scaled by a factor of 10. As the magnitudes of the muscle contraction forces are dependent on fibre stretch rather than mass, the animation equilibrium states were not affected. If other forces were included, such as gravity, the magnitudes of which are affected by mass, these would need to be scaled to prevent the mass scaling from greatly affecting the equilibrium states. Although mass scaling affects the transient response of dynamic simulations, producing a slower, more sluggish response, these effects can be reduced by scaling other parameters, such

as reducing damping, and playing back the animations at a faster speed. For more accurate animations, mass scaling could be applied selectively to the necessary elements, reducing the amount of non-physical mass introduced [53].

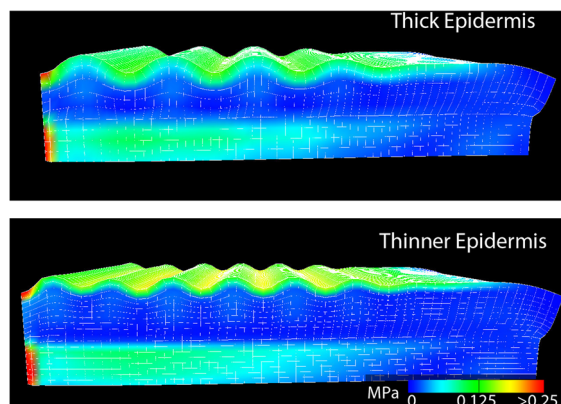
Time scaling was also used, whereby muscles were contracted at a faster speed and the animations played back slower, reducing the number of timesteps over which the animations were run. Although this can result in a more energetic transient response, these effects can be reduced by scaling other parameters, such as increasing damping. Using both mass and time scaling, damping parameters and the playback speed of animations were determined experimentally. As inertial effects are still relatively insignificant compared with muscle contraction forces when using such scaling parameters, the scaling techniques appeared to have little visual effect on the animations, particularly when played in real time.

## 7.2. Simulating the Epidermis

From Table I and Figure 7, it can be seen that, with the soft-tissue models, the epidermal layer (stratum corneum) was assigned a higher stiffness and appears thicker than in reality. Because of the low thickness of this layer in relation to element size, the dermis dominates the outer layer of elements. When the material properties of these skin layers are combined for these elements, this results in a stiffness much lower than that of the epidermis alone. The high stiffness is therefore necessary to produce a large enough difference between the stiffness of the two outer layers of elements, which is necessary to simulate wrinkles [3]. As a consequence, the outer layer of elements acts like a thick epidermal layer.

Better capturing of the epidermal thickness would probably produce a more realistic wrinkling behaviour. For example, Figure 12 shows animation results of variations of a soft-tissue block model, one with cube-shaped hexahedral elements and another with thinner, cuboid-shaped elements. The latter simulation produces finer, more realistic wrinkles, rather than larger bulges. However, as solid elements (e.g. hexahedra) have optimal accuracy and stability when regularly shaped, producing inaccurate and unstable results when highly distorted, the thickness of the epidermal layer is far too thin to be feasibly captured using such elements.

Shell elements can have a thickness that is much lower than the other element dimensions. Such elements are therefore suited to modelling thin objects, such as the epidermis, and have been successfully used in a previous work to model epidermal layers on detailed soft-tissue block models [3]. For the accuracy we desire, because of the low thickness of the epidermal layer, it will likely be sufficient to simply attach the mid-surface nodes of four-node quadrilateral shell elements to the outer faces of the outer



**Figure 12.** Stress visualisations of soft-tissue block models under muscle contraction. The models have identical element lengths and widths and element heights of 0.5 mm (left) and 0.25 mm (right). The top layer of elements containing the epidermal layer has the same material properties in both models.

skin elements. The shell elements would therefore share the same nodes as the solid elements, and not require any additional nodal tying constraints.

## 7.3. Material Behaviour

Although our soft-tissue animations use simple hyperelastic neo-Hookean materials, such materials normally only fit material behaviour with reasonable accuracy under moderate strains. As high strains are produced during skin wrinkling, a more accurate behaviour could be simulated by using, for example, Mooney–Rivlin or Ogden material models. Further skin behaviour could also be simulated, for example, by including anisotropic and viscoelastic material behaviours to model the preferred wrinkling direction parallel to Langer lines and skin properties such as hysteresis, stress relaxation, and creep. To model anisotropy in skin, like with muscles, each element overlapping an anisotropic skin layer would need to have one or multiple directions associated with that layer, for example, representing the directions of collagen fibres or Langer lines. The inclusion of biaxial natural tension with larger tension parallel to Langer lines would also help produce realistic anisotropic wrinkling effects. Such complex materials can be directly incorporated to the TLED FE method with little additional computational cost [54].

Modelling such complex material behaviour could also help to produce more realistic animations of different aged soft tissues. As well as changes in material properties with age, such as the stiffness and thickness of skin layers, effects such as the decrease in natural tension, increase in anisotropy, and change in collagen fibre orientation could be modelled, each of which has an effect on wrinkling behaviour [55]. The increased appearance of wrinkle lines with age could also be modelled by considering skin plasticity [44]. To produce a model of aged soft tissue with permanent ageing wrinkles, starting with an unaged model (such as that in Figure 7), a simulation of ageing involving numerous muscle contractions could be performed, with which, for example, the yield strength of the plasticity models could be decreased to simulate quick formation of the ageing wrinkles. From this, the initial deformation and material parameters for a series of different aged faces could be determined.

## 7.4. Simulation Models

In Figures 8 and 9, some artefacts can be seen on the forehead animations due to the inaccuracy of the surface mesh. The animations produced are highly dependent on the model and parameters; for example, the direction of movement and wrinkling behaviour, such as number and size of wrinkles, depend highly on muscle shape, size and direction, thickness and parameters of skin layers and soft-tissue structures, and constrained nodes. It is therefore likely that much more realistic models and animations could be produced by using models generated from

more accurate surface meshes (e.g. from computed tomography and MRI data), rather than manually created surface meshes, as well as more accurate materials and muscle parameters. Models could be created using medical data with our model creation process, although such data would first require processing and clean-up.

## 8. CONCLUSION

This work has presented a fully physically-based approach for efficiently producing realistic-looking animations of facial movement, including animation of expressive wrinkles. Focussing on the forehead, this involves using an optimised GPU-based non-linear TLED FE solver to simulate non-conforming multi-layered hexahedral soft-tissue models that can be generated using our model creation process. The simulation process includes an anatomical muscle contraction model that generates active and transversely isotropic passive muscle stresses, as well as procedures to constrain nodes. For example, nodes are able to slide over a non-conforming surface, enabling the sliding effect between superficial and deep soft tissues to be modelled.

Animation examples have demonstrated the ability of our process to produce animation of realistic large-scale and fine-scale soft-tissue behaviours, including wrinkles. Mass and time scaling are useful for such simulations to enable a much higher timestep to be used with little visual effect. The flexibility of our approach has enabled a variety of animations to be produced by using different parameters and simulation models, for example, to simulate different aged skins or even arbitrary multi-material soft-body structures.

However, various improvements could be made to the simulation process. Shell elements could be used to model the epidermis, and multi-resolution models could be used, for example, to enable a higher resolution to be employed along the outer skin surface where wrinkles are produced. More realistic animations would also likely be achieved by using more accurate models produced using medical data. Future work will focus on experimenting with models and parameters to produce more realistic animations of different-aged forehead movement, including incorporating more complex anisotropic viscoelastic material models into our system.

The next stages of the work will also involve evaluating our animation approach. For wrinkling of different-aged foreheads, as well as surveys where subjects select, for example, an age group and realism level for animations, qualitative evaluation could be assisted by using wrinkle classification scales, such as the Glogau classification of photoageing and the Fitzpatrick wrinkle scale to rate animated wrinkles, and comparing these with experimental data. Quantitative evaluation could also be performed by comparing simulated wrinkle measurements and trends to real wrinkling data [3], which could be estimated using medical data, digitised models of forehead wrinkles [14], or video.

## ACKNOWLEDGEMENT

This research was supported by the EPSRC.

## REFERENCES

1. Terzopoulos D, Waters K. Physically-based facial modeling, analysis, and animation. *Journal of Visualization and Computer Animation* 1990; **1**(2): 73–80.
2. Kähler K, Haber J, Seidel H-P. Geometry-based muscle modeling for facial animation, In *Proceedings of the GI*, Ottawa, Ontario, Canada, 2001; 37–46.
3. Flynn C, McCormack BAO. Finite element modelling of forearm skin wrinkling. *Skin Research and Technology* 2008; **14**(3): 261–269.
4. Kuwazuru O, Saothong J, Yoshikawa N. Mechanical approach to aging and wrinkling of human facial skin based on the multistage buckling theory. *Medical Engineering & Physics* 2008; **30**(4): 516–522.
5. Koch RM, Roth SHM, Gross MH, Zimmermann AP, Sailer HF. A framework for facial surgery simulation, In *Proceedings of the SCCG*, Budmerice, Slovakia, 2002; 33–42.
6. Zachow S, Hege H-C, Deuffhard P. Computer-assisted planning in cranio-maxillofacial surgery. *Journal of Computing and Information Technology* 2006; **14**(1): 53–64.
7. Fratarcangeli M. Physically-based synthesis of animatable face models, In *Proceedings of the VRIPHYS*, Pisa, Italy, 2005; 32–39.
8. Chen C, Prakash EC. Physically-based facial expression synthesizer with performance analysis and GPU-aided simulation, In *Proceedings of the CyberGames*, Perth, Australia, 2006; 171–176.
9. Fratarcangeli M. Position-based facial animation synthesis. *Computer Animation and Virtual Worlds* 2012; **23**(3–4): 457–466.
10. Waters K. A muscle model for animation three-dimensional facial expression, In *Proceedings of the SIGGRAPH*, Anaheim, California, USA, 1987; 17–24.
11. Zhang Y, Prakash EC, Sung E. A new physical model with multilayer architecture for facial expression animation using dynamic adaptive mesh. *IEEE Transactions on Visualization and Computer Graphics* 2004; **10**(3): 339–352.
12. Coull AD. A physically-based muscle and skin model for facial animation, *Ph.D. Thesis*, University of Glasgow, UK, 2005.
13. Reis CDG, Bagatelo H, Martino JM. Real-time simulation of wrinkles, In *Proceedings of the WSCG*, Pilsen, Czech Republic, 2008; 109–116.
14. Bando Y, Kuratate T, Nishita T. A simple method for modeling wrinkles on human skin, In *Proceedings of the PG*, Beijing, China, 2002; 166–175.

15. Müller M, Chentanez N. Wrinkle meshes, In *Proceedings of the SCA*, Madrid, Spain, 2010; 85–92.
16. Wu Y, Kalra P, Moccozet L, Magnenat-Thalmann N. Simulating wrinkles and skin aging. *Visual Computer* 1999; **15**(4): 183–198.
17. Zhang Y, Sim T, Tan CL. Simulating wrinkles in facial expressions on an anatomy-based face, In *Proceedings of the ICCS*, Atlanta, Georgia, USA, 2005; 207–215.
18. Hung A, Mithrathatne K, Sagar M, Hunter P. Multi-layer soft tissue continuum model: towards realistic simulation of facial expressions, In *Proceedings of the WASET*, Paris, France, 2009; 134–138.
19. Barbarino G, Jabareen M, Trzewik J, Mazza E. Physically-based finite element model of the face, In *Proceedings of the ISBMS*, London, UK, 2008; 1–10.
20. Mollemans W, Schutyser F, Nadjmi N, Suetens P. Very fast soft tissue predictions with mass tensor model for maxillofacial surgery planning systems, In *Proceedings of the CARS*, Berlin, Germany, 2005; 491–496.
21. Xu S, Liu XP, Zhang H, Hu L. A nonlinear viscoelastic tensor-mass visual model for surgery simulation. *IEEE Transactions on Instrumentation and Measurement* 2011; **60**(1): 14–20.
22. Beldie L, Walker B, Lu Y, Richmond S, Middleton J. Finite element modelling of maxillofacial surgery and facial expressions—a preliminary study. *International Journal of Medical Robotics and Computer Assisted Surgery* 2010; **6**(4): 422–430.
23. Taylor ZA, Cheng M, Ourselin S. High-speed nonlinear finite element analysis for surgical simulation using graphics processing units. *IEEE Transactions on Medical Imaging* 2008; **27**(5): 650–663.
24. Wu T, Mithrathatne K, Sagar M, Hunter PJ. Characterizing facial tissue sliding using ultrasonography, In *Proceedings of the WCB*, Singapore, 2010; 1566–1569.
25. Sifakis E, Neverov I, Fedkiw R. Automatic determination of facial muscle activations from sparse motion capture marker data. *ACM Transactions on Graphics* 2005; **24**(3): 417–425.
26. Warburton M, Maddock S. Creating animatable non-conforming hexahedral finite element facial soft-tissue models for GPU simulation, In *Proceedings of the WSCG*, Pilsen, Czech Republic, 2012; 317–325.
27. Warburton M, Maddock S. Creating finite element models of facial soft tissue, In *Proceedings of the WSCG*, Pilsen, Czech Republic, 2013; 215–224.
28. Lee Y, Terzopoulos D, Waters K. Realistic modeling for facial animation, In *Proceedings of the SIGGRAPH*, Los Angeles, California, USA, 1995; 55–62.
29. Nedel LP, Thalmann D. Real time muscle deformations using mass-spring systems, In *Proceedings of the CGI*, Hannover, Germany, 1998; 156–165.
30. Röhrle O, Pullan AJ. Three-dimensional finite element modelling of muscle forces during mastication. *Journal of Biomechanics* 2007; **40**(15): 3363–3372.
31. Lee S-H, Terzopoulos D. Heads up! Biomechanical modeling and neuromuscular control of the neck. *ACM Transactions on Graphics* 2006; **25**(3): 1188–1198.
32. Tang CY, Zhang G, Tsui CP. A 3D skeletal muscle model coupled with active contraction of muscle fibres and hyperelastic behaviour. *Journal of Biomechanics* 2009; **42**: 865–872.
33. Duysak A, Zhang JJ. Fast simulation of facial tissue deformations using mass-spring chain algorithm, In *Proceedings of the TPCG*, Canterbury, UK, 2005; 139–145.
34. Choe B, Lee H, Ko H-S. Performance-driven muscle-based facial animation. *Journal of Visualization and Computer Animation* 2001; **12**(2): 67–79.
35. Mithrathatne K, Hung A, Sagar M, Hunter PJ. An efficient heterogeneous continuum model to simulate active contraction of facial soft tissue structures, In *Proceedings of the WCB*, Singapore, 2010; 1024–1027.
36. Kim H, Jürgens P, Weber S, Nolte L-P, Reyes M. A new soft-tissue simulation strategy for cranio-maxillofacial surgery using facial muscle template model. *Progress in Biophysics & Molecular Biology* 2010; **103**(2): 284–291.
37. Kähler K, Haber J, Yamauchi H, Seidel H-P. Head shop: generating animated head models with anatomical structure, In *Proceedings of the SCA*, San Antonio, Texas, USA, 2002; 55–63.
38. Yang XS, Zhang JJ. Modelling and animating hand wrinkles, In *Proceedings of the ICCS*, Atlanta, Georgia, USA, 2005; 199–206.
39. Larboulette C, Cani M-P. Real-time dynamic wrinkles, In *Proceedings of the CGI*, Crete, Greece, 2004; 522–525.
40. Dutreux L, Meyer A, Bouakaz S. Real-time dynamic wrinkles of face for animated skinned mesh, In *Proceedings of the ISVC*, Las Vegas, Nevada, USA, 2009; 25–34.
41. Li M, Yin B, Kong D, Luo X. Modeling expressive wrinkles of face for animation, In *Proceedings of the ICIG*, Chengdu, Sichuan, China, 2007; 874–879.
42. Sanchez M. Techniques for performance-based, real-time facial animation, *Ph.D. Thesis*, The University of Sheffield, Sheffield, UK, 2006.
43. Bickel B, Botsch M, Angst R, Matusik W, Otaduy M, Pfister H, Gross M. Multi-scale capture of facial geometry and motion. *ACM Transactions on Graphics* 2007; **26**(3): 33-1–33-10.
44. Magnenat-Thalmann N, Kalra P, Lèvêque JL, Bazin R, Batisse D, Querleux B. A computational skin model:

- fold and wrinkle formation. *IEEE Transactions on Information Technology in Biomedicine* 2002; **4**(6): 317–323.
45. Bischoff JE. Reduced parameter formulation for incorporating fiber level viscoelasticity into tissue level biomechanical models. *Annals of Biomedical Engineering* 2006; **34**(7): 1164–1172.
  46. Miller K, Joldes G, Lance D, Wittek A. Total Lagrangian explicit dynamics finite element algorithm for computing soft tissue deformation. *Communications in Numerical Methods in Engineering* 2007; **23**(2): 121–134.
  47. Mosegaard J, Herborg P, Sørensen TS. A GPU accelerated spring mass system for surgical simulation. *Studies in Health Technology and Informatics* 2005; **111**: 342–348.
  48. Maciel A, Halic T, Lu Z, Nedel LP, De S. Using the PhysX engine for physics-based virtual surgery with force feedback. *International Journal of Medical Robotics and Computer Assisted Surgery* 2009; **5**(3): 341–353.
  49. Teran J, Sifakis E, Blemker SS, Ng-Thow-Hing V, Lau C, Fedkiw R. Creating and simulating skeletal muscle from the visible human data set. *IEEE Transactions on Visualization and Computer Graphics* 2005; **11**(3): 317–328.
  50. Joldes GR, Wittek A, Miller K. An efficient hourglass control implementation for the uniform strain hexahedron using the total Lagrangian formulation. *Communications in Numerical Methods in Engineering* 2008; **24**(11): 1315–1323.
  51. Avril Q, Gouranton V, Arnaldi B. Fast collision culling in large-scale environments using GPU mapping function, In *Proceedings of the EGPGV*, Cagliari, Italy, 2012; 71–80.
  52. Warburton M, Maddock S. GPU simulation of finite element facial soft-tissue models, In *Proceedings of the TPCG*, Bath, UK, 2013; 1–8.
  53. Olovsson L, Simonsson K, Unosson M. Selective mass scaling for explicit finite element analyses.

*International Journal for Numerical Methods in Engineering* 2005; **63**(10): 1436–1445.

54. Taylor ZA, Comas O, Cheng M, Passenger J, Hawkes DJ, Atkinson D, Ourselin S. On modelling of anisotropic viscoelasticity for soft tissue simulation: numerical solution and GPU execution. *Medical Image Analysis* 2009; **13**(2): 234–244.
55. Flynn C, McCormack BAO. Simulating the wrinkling and aging of skin with a multi-layer finite element model. *Journal of Biomechanics* 2010; **43**(3): 442–448.

## SUPPORTING INFORMATION

A supporting video showing the full animation process and various animation examples can be found in the online version of this article at the publisher's web site.

## AUTHORS' BIOGRAPHIES



**Mark Warburton** is a PhD student in Computer Science at the University of Sheffield. His research interests are in physically-based modelling and animation, facial animation, and graphics processing unit computing. Mark completed a BSc in Computer Science at the University of Sheffield in 2010, for which he was awarded a faculty medal. He lead-authored the best student paper at TPCG 2013.



**Steve Maddock** is a senior lecturer in computer science at the University of Sheffield. His research interests are in computer facial modelling and animation, human figure animation, procedural modelling, and surface deformation techniques. He received a PhD in computer science from the University of Sheffield in 1999. He is a member of ACM SIGGRAPH and Eurographics.