Physically-Based Facial Modelling and Animation including Wrinkles using the Finite Element Method on the GPU - 24 Month Report

Mark Anthony Warburton
Department of Computer Science
The University of Sheffield
October 2012

Abstract

Physically-based animation techniques enable more realistic and accurate animation to be created. Such approaches require two main stages: the creation and simulation of a model. Our model creation process, which automatically creates simulation-ready non-conforming hexahedral finite element (FE) models, and GPU FE solver are presented, along with examples of soft-tissue block and forehead simulations. However, there are currently issues when simulating wrinkling using the created non-conforming models. These issues are discussed, along with potential solutions, such as smoothing the simulation meshes.

1 Introduction

Using physically-based techniques for facial animation, the effects of muscle contractions can be propagated through a facial soft-tissue model to automatically deform the model in a realistic and anatomical manner. Simulation of physically-based models requires two main stages: creation of a suitable model, and simulation of the model through time using a physics-based technique, such as the mass-spring (MS) or finite element (FE) method.

Advancing the state of the art in facial animation for computer graphics applications, the aim of this project is to develop a fully physically-based technique using the accurate non-linear total Lagrangian explicit dynamic (TLED) formulation of the FE method for producing realistic-looking, natural animations of different aged human facial movement, including the simulation of both large and fine expressive and aging facial wrinkles. Due to the obvious wrinkles that can be produced in the forehead region, and the simpler anatomy of this area in comparison to other areas of the face, such as around the mouth, the main focus of the project is on simulating forehead movement and wrinkles, rather than the entire face.

While high-quality conforming simulation models are usually required for high accuracy applications, non-conforming models can enable more efficient production of stable, realistic-looking animations for computer graphics applications, while being much easier to create, and are therefore being used for this project. Further optimisation of the simulation of such models can also be achieved as CUDA is being used to solve the TLED FE equations in parallel on the GPU, increasing the performance gained by parallelising the simulations.

Over the past 12 months, developing on current techniques for producing non-conforming hexahedral meshes with bound surface meshes (e.g. [DGW11]), the main focus of this project has been to develop an automatic process to easily construct animatable non-conforming hexahedral FE simulation models with cuboid-shaped elements. This includes automatic computation of element material types, boundary conditions and muscle properties. While the focus has been on creating facial soft-tissue models (the soft tissue between the skull and outer skin surface, as show by Figure 1.1), the process can be used to create any multi-layered model from any surface meshes. The developed CUDA TLED FE system has also been optimised to exploit the computational advantages offered by using such non-conforming hexahedral simulation models [WM12]. However, some problems have been encountered when trying to simulate fine details, such as wrinkles, using non-conforming cuboid elements that contain sharp steps, which could potentially be solved, for example, by smoothing the model before simulation [AF07].
2 Further Initial Experiments

2.1 Comparison of Conforming and Non-Conforming Simulation Meshes

For our work, non-conforming hexahedral FE simulation meshes with cuboid-shaped elements have been used. However, simulation meshes that conform to surface meshes can also be used. Table 2.1 summarises the creation and use of conforming and non-conforming hexahedral meshes with a CUDA TLED FE system. As well as model creation and stability advantages, various performance advantages can be gained by using the non-conforming approach with a CUDA TLED FE system:

- Only a single set of the various element values, such as a single set of shape function derivatives (12 values) and one hourglass stiffness matrix (64 values) when hourglass control is used, needs to be stored, rather than a set for each element. While greatly reducing memory usage, this also enables more efficient memory accesses as the same set of values are accessed by each thread, reducing the required number of slow global GPU memory accesses.

- Elements and element nodes can be easily efficiently numbered, which not only means that only 4 node indices per element need to be stored, from which the other 4 indices can be de-

---

Table 2.1: A summary of the differences between using a conforming and non-conforming hexahedral (cuboid-element) simulation mesh with a CUDA TLED FE system.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Conforming</th>
<th>Non-Conforming</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volumetric Model Generation</td>
<td>Complex, involving manual input</td>
<td>Simple and automatic</td>
</tr>
<tr>
<td>Simulation Speed</td>
<td>Difficult to achieve speed-ups</td>
<td>More efficient memory accesses and global memory coalescing can be achieved for increased speed</td>
</tr>
<tr>
<td>Memory Requirements</td>
<td>High, especially with large models</td>
<td>Low as some values are the same for each element</td>
</tr>
<tr>
<td>Stability</td>
<td>Badly-shaped or tiny elements can greatly reduce stability</td>
<td>All elements are well-shaped for better stability</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Mesh more closely matches the object being simulated</td>
<td>Less accurate approximation of the object being simulated may lead to less accurate results</td>
</tr>
</tbody>
</table>

In the following sections, a comparison between using conforming and non-conforming simulation meshes, particularly for GPU FE simulations, is given, followed by an overview of the current multi-layer non-conforming hexahedral model creation process, and CUDA TLED FE simulation system. Various simulation examples are also shown, including wrinkling simulation on a soft-tissue-block model, and a simulation of frontalis muscle contraction to raise the eyebrows on a facial model. The problems with the current approach to facial wrinkle simulation are then described, along with potential solutions. The final sections present a plan for the final PhD thesis, as well as a plan for the remaining project work to be completed.
Table 2.2: Results of a comparison between using a conforming and a non-conforming hexahedral (cuboid-element) simulation mesh of a spherical object, constrained by its upper nodes and acting under gravity, with a Neo-Hookean material model (with a default Young’s modulus, \( E \), of 3 kPa and Poisson ratio, \( v \), of 0.4). Simulations were performed on a NVIDIA GeForce GTX 460 1GB GPU.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Conforming</th>
<th>Non-Conforming</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodes</td>
<td>2736</td>
<td>2849</td>
</tr>
<tr>
<td>Elements</td>
<td>2197</td>
<td>2176</td>
</tr>
<tr>
<td>Minimum Element Length (mm)</td>
<td>&lt;12.5</td>
<td>12.5</td>
</tr>
<tr>
<td>Time-Step Solution Time (ms)</td>
<td>0.48</td>
<td>0.246</td>
</tr>
</tbody>
</table>

Stability with a 0.1ms timestep:

- Maximum \( E \) (kPa): 3 kPa vs. 10 kPa
- Maximum \( v \): 0.4 vs. 0.47

Figure 3.1: Voxel element samples assigned a particular material, and a fibre direction for each overlapping muscle, which are used to calculate the overall element material properties and fibre directions. Note this is a 2D illustration of a 3D process.

Experiments have demonstrated these performance and stability improvements. As shown by Table 2.2, using a non-conforming simulation mesh with a bound higher resolution surface mesh has led to performance increases of almost 2x compared to using a conforming simulation mesh with roughly the same number of nodes and elements. Also, stable simulations for FE simulations were able to be performed using a considerably higher Young’s modulus and Poisson ratio with a non-conforming mesh, which is necessary for simulating, for example, the stiff properties of the epidermal skin layer, and incompressible soft-tissue material. On the downside, depending on element size, accuracy is likely to be reduced using a non-conforming simulation mesh.

3 Model Creation

The model creation process described in this section extends upon our earlier version of the process [WM12]. The current process consists of five main stages:

1. Voxelising the surface mesh
2. Computing element material properties
3. Computing muscle fibre directions
4. Determining nodes that are rigid or bound by a surface mesh (sliding nodes)
5. Binding the surface meshes to the simulation model

The surface mesh can contain various surfaces, and elements are separated depending on the user-defined collection of surfaces (volumes) by which they are enclosed. For example, with a facial
3 Model Creation

Figure 3.2: Rigid nodes on a single side of a 3D simulation model aligned with the global axes. All nodes are therefore on the outside of the model, although some nodes are too far away from the rigid surface, while the angle between the normal of some others (marked with an X) and the closest point on the rigid surface is greater than $\pi/2$.

mesh, there may be a volume for the skin and connective tissue (between the outer skin surface and the skull), and a volume for each muscle. The voxelisation process uses a number of samples per voxel to approximate the proportion overlap between each voxel and the overlapping volumes, as shown by Figure 3.1. Voxels can either be regularly (cubes) or irregularly shaped (cuboids).

Using the samples, element constitutive models, mass densities and muscle stresses are weighted according to the proportion of volumes overlapping the element. As well as material properties, muscle fibre directions are also calculated at each element sample for each muscle that overlaps the element. Within the skin and connective tissue volume, constant thickness layers with different material properties can also be automatically created by simply testing the distance from the outer skin surface of samples associated with this volume to determine which layer the sample is contained within.

Sample fibre directions are calculated using NURBS volumes, which are created by shrinking NURBS surfaces to their central curve [D. 01]. Taking the partial derivative of the function defining the NURBS volume, $V$ with respect to the material (parametric) coordinate along the length of the muscle, $a$, produces an implicit fibre field, $d(x)$ [TSB+05]:

$$d(x) = \frac{\partial V(V^{-1}(x))}{\partial a} \left/ \left\| \frac{\partial V(V^{-1}(x))}{\partial a} \right\| \right.$$  \hspace{1cm} (3.1)

where $x$ is the sample spatial coordinate.

As the NURBS volume function requires material rather than spatial coordinates, the material coordinates of sample points, $V^{-1}(x)$, are estimated using the Newton-Raphson root finding method.

As described in Section 4.1, using our simulation system, it is possible to set nodes as rigid or sliding (bound by a surface), whereby rigid nodes remain fixed throughout a simulation, and sliding nodes remain a particular distance from the surface mesh they are bound to (e.g. to enable soft tissue to slide over the skull surface). To identify such nodes given a collection of rigid and sliding surfaces, for each internal rigid and sliding surface, the closest point on the surface to each node is found, and, if this is located within a particular bounding box of the node, that node is set to rigid or sliding. For external surfaces, a similar approach is taken, except the node must be on the outside of the model, and the angle between the node normal and the surface point normal must be less than $\pi/2$ (i.e. these vectors must be pointing in a similar direction), as shown by Figure 3.2.

For the final stage of the model creation process, as with Dick et al.’s simulation approach [DGW11], the vertices of the surface meshes can simply be bound to and animated with elements of the FE mesh using trilinear interpolation and extrapolation. Figure 3.3 shows two examples of
Figure 3.3: Soft-tissue block (top) and facial (middle and bottom) simulation models. Red spheres represent rigid (fixed) nodes, while green elements overlap a muscle. The top and middle left images show views from behind the models to show the rigid nodes and muscle elements.
simulation models generated using our model creation process - one of a soft-tissue block model with a relatively flat cylindrical muscle, and one of a simple facial model with the frontalis muscle.

The facial surface, including the inner surface used as an approximation of the skull, was generated using FaceGen\(^1\), and the NURBS surfaces for the frontalis muscle were modelled using Rhino 3D\(^2\). As the face model only includes the frontalis muscle, only the relevant section of the face containing this muscle was voxelised. Although, in reality, the frontalis muscle fibres blend with the galea aponeurotica that continues around the top and back of the head, and therefore have no attachment to the skull, the galea aponeurotica acts like an anchor for the frontalis on the top of the head, causing the majority of soft-tissue movement produced by contraction of the frontalis to be in the forehead region. With the developed facial model, the frontalis simply blends with the skull around the top of the head, and the intersection of the frontalis and the skull is treated as a rigid surface to model the anchoring effect of the galea aponeurotica. The remainder of the skull surface is treated as a sliding surface. Due to the simplicity of the model, it was also necessary to set the nodes along the tops of the eyes as rigid to further constrain movement around this area and prevent large gaps forming above the eyes during simulations.

The model creation process generally produced desirable results, for example, with the desired facial model nodes set to rigid or sliding on a complex surface, and muscle fibre directions calculated with reasonably good accuracy for each element that overlaps a muscle. Simulation examples using these models are presented in Section 5.

4 Model Simulation

4.1 CUDA TLED FE System

Work has been done developing the non-linear TLED FE solver on the GPU, introduced in the 12 month report [War11]. As well as 4-node tetrahedral elements and the Neo-Hookean material model, this now also supports reduced-integration 8-node standard hexahedral and cuboid elements (used with the non-conforming models). Compared with 4-node tetrahedra, such elements offer stability, performance and accuracy advantages, particularly when modelling non-linear incompressible materials [TCC\(+\)09, fLhLfT11]. The system has also been optimised to exploit the various advantages offered when simulating cuboid elements on the GPU (described in Section 2.1).

In order 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 occur with under-integrated 8-node hexahedral elements, a hourglass control method has been implemented. The implemented technique is a stiffness-based method, which adds an artificial stiffness to elements to constrain the different hourglass modes based on the element stiffness and a user-defined hourglass stiffness parameter [JWM08]. Using this method, an element hourglass force matrix, \( tF^H(m) \), is calculated:

\[
\begin{align*}
\{F^H(m)\}^T &= 0H(m)\{U(m)\} \\
0H(m) &= \kappa \cdot k_{Max} \cdot 0\gamma(m) \cdot 0\gamma(m)^T
\end{align*}
\]

where \( 0H(m) \) is the hourglass matrix, \( tU(m) \) is a matrix of nodal displacements of element \( m \), \( \kappa \) is a specified stiffness parameter, \( k_{Max}^{(m)} \) is the maximum element stiffness, and \( 0\gamma(m) \) 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 can be simply added to the relevant nodal internal forces. In the above equations, the hourglass matrix can be pre-computed, and, as a general rule, values of \( \kappa \) should be used such that the total hourglass energy doesn’t exceed 10% of the internal energy to prevent materials appearing overly stiff.

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, the muscle model uses the fibre direction to compute an additional passive stress component to model this transverse isotropy.

\( ^1 \)http://facegen.com/
\( ^2 \)http://www.rhino3d.com/
Section 4 Model Simulation

![Diagram of sliding nodes moving along a sliding surface with force direction and direction to closest point on sliding surface indicated.](image)

Figure 4.1: Sliding nodes moving along a sliding surface. The displacement of sliding nodes before (image 2) and after (image 3) the additional sliding displacement has been considered is shown. Note this is a 2D illustration of a 3D process.

For a particular muscle, the additional second Piola-Kirchhoff stress produced by the muscle, \( {\sigma}^{(\text{mus})} \), for element \( m \) is computed by adding the active stress, \( {\sigma}^{(\text{act})} \), and additional passive stress, \( {\sigma}^{(\text{pas})} \):

\[
{\sigma}^{(\text{mus})} = {\sigma}^{(\text{act})} + {\sigma}^{(\text{pas})}
\]

where \( J \) is the Jacobian of the deformation gradient, \( \sigma^{(\text{act})} \) and \( \sigma^{(\text{pas})} \) are the active and passive muscle Cauchy stress references respectively, both weighted by the element overlap with the muscle.

With the current system, \( f^{(\text{act})} \) and \( f^{(\text{pas})} \) created by Rührle et al. were used [RP07].

To constrain models during simulation, it is possible to set nodes as rigid or sliding (bound by a surface). Rigid nodes are simply fixed with zero displacement throughout the simulations, whereas sliding nodes maintain a fixed distance away from the non-conforming sliding surface they are associated with. The displacement of a sliding node, \( q \), is updated after computation of a user-specified number of timesteps to move this node to a position \( q_{s} \) (the fixed distance) away from the sliding surface, \( s \). This additional displacement, \( u^{(q,s)} \), is calculated as:

\[
u^{(q,s)} = -b^{(q,s)} \cdot \left( d^{(q,s)} + b^{(q,s)} \cdot \| (p^{(q)} - c^{(q,s)}) \| \right) \cdot \frac{(p^{(q)} - c^{(q,s)})}{\| (p^{(q)} - c^{(q,s)}) \|}
\]

\[
b^{(q,s)} = \begin{cases} 1 & \text{for } (p^{(q)} - c^{(q,s)}) \cdot n^{(q,s)} < 0 \\ -1 & \text{otherwise} \end{cases}
\]

where \( p^{(q)} \) is the position of node \( q \), \( c^{(q,s)} \) is the closest point on surface \( s \) to node \( q \), and \( n^{(q,s)} \) is the surface normal at this point.

On a facial model, rigid nodes can be used to model the effect of restraining ligaments, whereas sliding nodes can model the sliding effect that occurs between the superficial and deep facial soft-tissue layers [WMSH10].

To increase computational performance, a GPU semi-brute broad-phase collision detection algorithm with uniform-grid spatial subdivision has been implemented [AGA12] to prune the number of polygons on sliding surfaces to be tested when finding the closest position for each sliding node. As shown by Figure 4.1, sliding nodes are able to follow the shape of (slide around) a surface, even though the elements don’t necessarily conform to that surface.

In related work, a penalty method has been used to add artificial forces to push nodes away from a surface that they penetrating [LTW95, HWHM11]. While potentially more physically accurate, such a method is more complex, and could cause instabilities depending on the size of the forces generated and the material properties of the elements. Such a penalty method, however,
5 Simulation Examples

5.1 Simulation Setup

For each of the simulations described in the rest of the report, unless stated otherwise, the material properties in Table 5.1 have been used, with a muscle active and passive stress reference of 50 MPa for each muscle. As the deep layers are tough and fairly rigid, these are not considered in any of the models produced so far, and the superficial layers simply slide over the skull or bone surface.

While the actual density of skin, connective tissue and muscle is approximately 1100 kg/m$^3$ [Fly07], stability issues arose when using such a low density, even when an extremely small timestep (as small as 1 µs) was used. A common technique to counter such issues with explicit simulations is to use mass or load scaling [OSU05]. Since the critical timestep for simulating a material using an explicit simulation, provided all other material properties remain constant, is roughly proportional...
Simulation Examples

<table>
<thead>
<tr>
<th>Layer</th>
<th>Density (kg/m³)</th>
<th>Young's Modulus (MPa)</th>
<th>Poisson Ratio</th>
<th>Thickness (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stratum Corneum</td>
<td>10,000</td>
<td>6</td>
<td>0.49</td>
<td>0.02</td>
</tr>
<tr>
<td>Dermis</td>
<td>10,000</td>
<td>0.0814</td>
<td>0.49</td>
<td>1.5</td>
</tr>
<tr>
<td>Hypodermis</td>
<td>10,000</td>
<td>0.034</td>
<td>0.49</td>
<td>Remainder</td>
</tr>
<tr>
<td>Muscle</td>
<td>10,000</td>
<td>0.5</td>
<td>0.49</td>
<td>Variable</td>
</tr>
</tbody>
</table>

Table 5.1: The Neo-Hookean material properties used for each of layers in the simulation examples. The hypodermis layer surrounds muscles and fills the remainder of the volume between the outer skin surface and the skull.

to the material density [JB96], by increasing the density of the materials, the critical timestep is increased.

By scaling the densities of the materials by a factor of 10, stable simulations can usually be run on models with skin and soft-tissue-like properties using a timestep of between 0.01 ms and 0.1 ms depending on element size. However, one potential problem with mass scaling is that the transient response of the dynamic simulations can be affected. Although further experimentation is required to determine the effect mass scaling has on the soft-tissue simulations (as current work has mainly focussed on examining the equilibrium states). At least with simulations of slower movements, however, the inertial forces should be relatively insignificant [SNF05], and it is expected that the transient response shouldn’t be affected too much in these cases.

5.2 Soft-Tissue Block Simulations

The initial, simpler version of the model creation process determined whether a voxel is inside a volume, and consequently its material type, using a single sample at the voxel centre, and the skin and hypodermis are treated as a single rather than multi-layer material. With such models, all nodes on a rigid surface, such as the skull, are also set as fixed with zero displacement throughout the simulations, and muscles contract towards a single centre of contraction. Figure 5.1 shows the equilibrium states of simple simulations involving a block of soft tissue being influenced by the contraction of a single muscle using simulations models of different complexity and anatomical accuracy. The most anatomically accurate model was created using the full model creation process (see Section 3 and Figure 3.3). As expected, best results are achieved by using the most anatomically accurate model with three layers of skin and connective tissue, accurate muscle fibre directions, and nodes that are able to slide across the skull/bone.

Modelling the skin and connective tissue as a layered structure seems to have the largest effect. As shown by Figure 5.1, no wrinkling is produced when this region is modelled using a single material, even if the material properties are calculated as an average of those used with the layered models. Due to this, no wrinkles are produced when a single sample at the centre of each voxel is used to calculate element properties, as there are no samples close enough to the outer skin surface to capture the material properties of the epidermis. This result is consistent with previous findings - projects that have simulated soft-tissue blocks containing various skin layers have been able to simulate wrinkles (e.g. [Fly07, HMSH09]), whereas simulations of other high-resolution FE models, such as the facial model by Sifakis et al., have produced large soft-tissue-like deformations without modelling multiple skin layers, although wrinkling is not present in these simulations [SNF05].

When the muscle contraction direction is simply towards a central contraction point, rather than along a complex fibre field, wrinkles still form, although some strange wrinkling effects seem to occur, particularly towards the muscle origin. On the other hand, when the bone is treated as a rigid rather than a sliding surface, wrinkles similar to those produced by the most anatomical model start to form, although the restricted movement prevents these wrinkles from forming fully.

5.3 Facial Simulations

As well as wrinkle simulation using a soft-tissue-block model, simulations using the facial model (Figure 3.3 in Section 3) have also been produced, as shown by Figure 5.2. The gross forehead movement, including eyebrow raising, has been successfully simulated. Some artefacts appear around the top of the nose at the edge of the simulation model. These are caused by deformations
Figure 5.1: The equilibrium states of soft-tissue block simulations under influence of a single muscle. The models for each simulation has different anatomical accuracy, with the most accurate being created using the full model creation process described in Section 3.
occurring around this area, causing polygons that lie across the edge, with some vertices bound to the simulation model and some static vertices, to be stretched. This could be rectified by further constraining the model, or extending it model to simulate more of the face. More accurate simulations could probably also be achieved by using more anatomically accurate skull and muscle models. However, the main issue with the simulations is that finer details, such as wrinkles, don’t appear, even when high resolution simulation models and surface meshes are used.

6 Current Issues

After research and experimenting with different simple soft-tissue-block models, it seems that the sharp steps produced on a non-conforming simulation model have a large effect on the ability to simulate wrinkles, even with high resolution models. Figure 6.1 shows a flat simple soft-tissue-block model that is aligned with the global axes. As the model size is $15 \times 30 \times 4.5\,mm$ and the regular elements have length $0.75\,mm$, the simulation model in this case is essentially conforming, like those produced in Figure 5.1 in Section 5.2. The layers and material properties stated in Table 5.1 in Section 5.1 have been used with Neo-Hookean material models. Nodes on the rear surface are rigid, and the bottom and side surfaces are sliding surfaces. Under muscle contraction, as expected, wrinkles are produced. Similar wrinkling results are also produced when a non-conforming flat model is used, although a thicker epidermal layer ($0.2\,mm$) was used to ensure at least one sample was contained within this layer during the model creation process. Figure 6.1 also shows a conforming model that is curved rather than flat, in which case the muscle contracts around the bottom curved sliding surface. Wrinkles are also produced in this simulation using a conforming curved simulation model.

However, as shown by Figure 6.2, wrinkles aren’t produced when using the curved skin-block mesh with a non-conforming simulation model, or when using a flat skin-block mesh that is angled such that the outer skin surface isn’t aligned with a global coordinate plane. Looking at the stresses produced, the most likely reason for this is due to the sampling resolution during model creation not being high enough to capture the properties of the thin epidermal layer passing through the middle of the elements, which, as previously mentioned, can also happen when using a non-conforming flat model. As shown in Section 5.2, if multiple skin layers aren’t modelled, wrinkles aren’t produced during simulations. Further experiments will need to be performed to determine if this is the problem, in which case simple solutions include increasing the sampling
Figure 6.1: Wrinkles simulated using a conforming flat, non-conforming flat, and conforming curved soft-tissue block model. In this model, the muscle has the same length and width as the skin and connective tissue, and a height of roughly 2\text{mm} (whereas the model height is 4.5\text{mm}). Simulation stresses are displayed, where blue indicates low, and red indicates high stress.
resolution or the thickness of the epidermal layer. Alternatively, the outer \( n \) layers of samples of the outer elements could be automatically set as epidermal samples.

Another problem with the non-conforming curved meshes that can be seen in Figure 6.2 is that artefacts and high stresses tend to be produced at locations where there is a sharp step in the simulation model. One possible solution would be to remove these sharp steps by smoothing the simulation model [AF07], and this technique is currently being implemented and tested. Although the simulation model still won’t conform exactly to the surface meshes, the sharp steps will be removed, which should enable more accurate stresses to be produced. Firstly, just the outer elements of the model should be smoothed, as the stresses produced at locations where elements cross the internal muscle boundaries seemed relatively smooth. The inner, non-conforming part of the model will therefore still possess the advantages of non-conforming simulation meshes, whereas the smoothed outer elements will ensure the simulation model conforms more closely to the overall shape of the object being modelled. It will be necessary to ensure enough outer layers of elements are smoothed to prevent highly distorted elements being produced. The material composition of these smoothed elements may also need to be altered since they will now overlap slightly different parts of the model, and possibly, for example, overlap a different proportion of the epidermal and dermal skin layers.

It may also be possible to approximate the model being simulated with a simpler-shaped mesh for which a conforming simulation model can be easily produced. In this case, the simulation model wouldn’t conform to the surface meshes of the object being simulated, but, like with the smoothing approach, there would be no sharp steps in the simulation model. Another possible solution could be to perform the simulation of wrinkles on a flat soft-tissue model, and map these wrinkles onto the simulation already produced by the non-conforming simulation model. This would probably be the least ideal solution, however; for example, the wrinkles won’t be produced by the simulation model of the actual object being simulated, making them less anatomically accurate, and there are likely to be issues when mapping wrinkles produced on a flat surface to a surface that is slightly curved.

As mentioned in Section 5.1, simulation stability issues have also been experienced. Although such issues are now much less of a problem, if further stability issues arise, other potential solutions
to increasing the critical timestep include using reduced order models [Tay11], or using a quasistatic rather than dynamic finite element formulation, although quasistatic analyses will only be suitable if the dynamic response of the simulations is insignificant. While an unconditionally stable implicit analysis would solve the stability issue, such analyses are more computationally intensive than explicit analyses. They are also difficult to parallelise, and would require a fair amount of work to implement.

7 Remaining Work

7.1 Further Implementation

As well as mesh smoothing and more accurate material property computation for element samples with non-conforming models, further functionality could also be implemented to help simulate realistic-looking forehead wrinkles without artefacts. For example, 2D shell elements could be used to model the stiff and extremely thin outer epidermal layer, rather trying to combine its material properties with the average material properties of thicker 3D elements.

By converting the surface meshes to meshes containing quad rather than triangle primitives, these quads could be used as shell elements to create a conforming models (it would be necessary to ensure there are no badly-shaped quads in the model, even in complex areas such as around the eyes). In this case, a mapping would be needed to propagate the stresses and forces between elements and nodes of the non-conforming 3D simulation model and those of the conforming 2D shell element simulation model, which probably wouldn’t be a trivial task. Using such an approach, gross movement could be efficiently simulated on the GPU using the current non-conforming hexahedral simulation models, which could be propagated to the conforming shell element simulation models to hopefully simulate wrinkling.

No known research has experimented with creating a mapping between two FE simulation models in this way, although simulations have been performed using soft-tissue blocks constructed of hexahedral elements with an outer layer of shell elements, in which case no mapping is required as the hexahedral and shell elements share the same nodes (e.g. [Fly07, WY09]). Depending on time constraints and other potential, more appropriate solutions found with further research, this shell element functionality may not be implemented.

Other functionality to be implemented includes additional material behaviour. Due to the flexible system design, integrating new material models into the system, such as the non-linear hyperelastic Mooney-Rivlin or Ogden material models, which are more general and usually more accurate than the Neo-Hookean model at higher strains. For soft-tissue and facial simulations, the Mooney-Rivlin model is popular (e.g. [SNF05, HMSH09]), and parameters for soft-tissue are therefore easily accessible, whereas parameters for the Ogden material model can be difficult to obtain without specific data collected from laboratory experiments; therefore, the Mooney-Rivlin model should be implemented first, followed by the Ogden model if necessary and if time allows.

The technique of assigning a preferred directions to element materials could also potentially be used with materials other than those used with muscles, such as skin layers and the hypodermis, making them transversely isotropic or anisotropic, and hence increasing the anatomical accuracy of the models. Preferred soft-tissue directions could be set at key points and interpolated to set preferred directions for samples and elements.

7.2 Evaluation

Various evaluation techniques were identified in the 12 month report [War11], and some of these will be used to evaluate the final system. Although the aim of the project is mainly to produce animations of realistic-looking wrinkles, to fully evaluate the system, both qualitative and some quantitative evaluation should be performed. Ethical clearance will be required to create and use any captured videos and data, and questionnaires for evaluation.

For qualitative evaluation, comparisons with real soft-tissue and forehead movement could be performed, and questionnaires could be used to get feedback from people with different amounts of computer graphics knowledge. Videos of people making particular forehead movements could be used along with videos of the same expressions being simulated with our simulation system. Provided the videos are filtered to make them all look like 3D graphics models, and each video and simulation is filtered or rendered to contain the same level of detail with a simple background, forehead movement and wrinkling produced by the system could be compared to that of real
people. In a survey, subjects could be asked to identify the simulation from the videos. For these survey results to be valid, it is thought that between 5 and 10 faces will be required for each comparison. Either the full face (provided only the forehead region moves in each video) or just the forehead region could be used in the videos.

If different wrinkling effects can be successfully simulated for different aged soft tissue, wrinkling simulations of different aged foreheads could be compared. In surveys, subjects could be asked to identify an age range for a given simulation, which will also reveal whether different wrinkling effects can make a model appear older using the system. The Glogau classification of photoaging [Bra06] and Fitzpatrick wrinkle scale [Der10] could also be used to rate the simulated wrinkles, and these could be compared to experimental results of wrinkle data. The ratings could be performed by a single or group of evaluators with different levels of expertise.

For quantitative evaluation, generic wrinkling data from experiments on different aged skin and faces is available [ANK+02]. Alternatively, wrinkling data could be estimated from a video capture session with multiple different aged subjects, or more accurate measures could also be used to obtain such data. Similar wrinkling data could then be calculated during simulations, and compared against the captured data. The accuracy of the FE solver could also be evaluated by comparing simple model simulations (e.g. of a simple soft-tissue block) with those simulated using an existing FE package.

8 Thesis Plan

The following sections briefly overview the planned chapters of the final PhD thesis. Where necessary, the practical work completed so far, and the remaining practical work to be completed is also described for each chapter. Chapters 2 and 3 constitute the literature review, which will mainly extend upon the literature review and initial experiments presented in the 12 Month Report [War11], whereas Chapters 4 to 7 detail the main contribution of the PhD.

8.1 Introduction

A general introduction of physically-based soft-tissue simulation systems will be presented, highlighting the two main categories - those used to produce realistic looking animations for computer graphics, and those with high physical accuracy used for studying soft-tissue behaviour. An introduction to the creation of types of physics-based models, such as conforming and non-conforming models, will also be presented, along with an introduction to the various physically-based simulation techniques, such as the mass-spring (MS) and finite element (FE) methods. Before giving an overview of the thesis chapters, the contributions of the project will then be stated, including:

- A technique for producing highly realistic and accurate physics-based facial animations efficiently on the GPU, including physics-based animation of large and fine wrinkles, advancing the current state of the art in computer graphics facial animation
- A method to automatically create simulation-ready physics-based models for producing the animations

8.2 Anatomy of the Human Head

Similar to the 12 Month Report, relevant anatomical information of the human head will be explained, including the anatomy of:

- Skin layers
- The skull
- Muscles
- The anatomy and properties of different types of wrinkles
8.3 Physically-Based Modelling and Animation Approaches

A further introduction to different physically-based simulation techniques, including different formulations of the FE method, will be presented. The more advanced mathematical theory will be explained in a section in the appendix. As in the 12 Month Report, a review of current physics-based techniques for modelling and simulating different objects and parts of the face will be presented, including:

- The skull
- Muscles
- The face, and blocks of skin or deformable soft tissue
- Wrinkles
- Generic deformable soft-bodies

There will also be a review of methods to create conforming and non-conforming physics-based simulation models with different finite element types, such as 4-node tetrahedra and 8-node hexahedra, including a description of the differences between these element types.

To further justify the use of a GPU-based non-linear total Lagrangian explicit dynamic (TLED) FE method, a more detailed comparison between physically-based simulation techniques, including details of the initial comparison experiments detailed in the 12 Month Report, will be presented. To further justify the use of a non-conforming, as opposed to a conforming, hexahedral simulation mesh, a comparison between using these two mesh types with a GPU-based TLED FE simulation system will also be presented, including results of simple example simulations. This comparison has also been reported in our conference paper [WM12].

8.4 Model Creation

The multi-layer non-conforming hexahedral model creation process, briefly presented in Section 3 and used to produce the models for this project, will be described, an earlier version of which has been detailed in our conference paper [WM12]. This will include a description of the following components:

- Voxelising the surface meshes
- Computing element material properties and muscle contraction directions using samples within voxels
- Creating NURBS volumes by shrinking NURBS surfaces to a central curve, and using these to determine the muscle fibre direction of a voxel sample
- Determining which simulation nodes are rigid, and which are bound by an impenetrable surface, such as the skull
- Binding a surface mesh to the non-conforming simulation mesh

The entire model creation process overviewed above has been completed, and there are no plans to further modify any of the above stages.

The problems with using a voxel-based simulation mesh containing sharp steps will also be referred to in this section, as well as techniques for countering these. This will include a description of the mesh smoothing process currently being implemented (see Section 6).

8.5 Model Simulation

Details of the GPU FE simulation system will be presented, including:

- The implementation and parallelisation of the TLED FE method, including some of the main details of the TLED FE formulation, and the relevant aspects of CUDA, with more advanced descriptions in the appendix
- The implementation of the material and element types, including the implementation of hourglass control for under-integrated 4-node hexahedra, and the optimisation to improve computational performance when using non-conforming cuboid elements
9 Conclusion and Project Plan

- The method used to generate active muscle stresses for muscle contraction, including computing contraction sizes and directions
- The algorithm developed to prevent nodes penetrating impenetrable surfaces, including the GPU collision detection algorithm implemented to optimise this process
- The overall flexible model creation and simulation system design that has enabled easy integration of new material and element types, and has allowed experiments to be easily performed using the constructed simulation models with various different model parameters.

The overall GPU simulation system has been completed, and includes implementations for tetrahedral, standard hexahedral and cuboid elements, and the Neo-Hookean material model. Some extensions may need to be made for better wrinkle simulation, such as the use of shell elements and more advanced material types (see Section 7.1).

8.6 Wrinkle Simulation

As a first stage to producing wrinkles, a simple geometric approach has been developed that can add a layer of geometric wrinkles on top of a model that is being simulated in a physically-based manner, such as an FE simulation of soft tissue, by analysing displacements of vertices in relation to muscle contraction directions. Details of this algorithm will be presented here, although it currently needs updating to work with the current simulation system. Any extension to the simulation system used to simulate wrinkles will also be detailed in this section, although no such extensions have currently been implemented.

8.7 Simulation Examples and Evaluation

Results (including performance, stability and output) and a comparison and analysis of the simulation system to other facial and wrinkle animation systems will be presented. The evaluation of the simulation system will also be presented. This will firstly include a description of the evaluation techniques used (both qualitative and quantitative), and how data for evaluation was collected. Analysis of the collected evaluation data, and a comparison with other systems will then be presented.

8.8 Conclusion

The conclusion will round up and summarise the thesis and contributed work, including a description of limitations and further work.

8.9 Appendix

The appendix, similar to the 12 Month Report, will contain sections on the following topics:

- Mathematical Theory - a detailed description and mathematical formulations of the MS and TLED FE methods
- GPU Computing - details of GPU Computing techniques, and details of CUDA that are relevant to the implemented GPU simulation system

9 Conclusion and Project Plan

Realistic and accurate physically-based animations require two main stages: the creation of a suitable model, and the simulation of the model through time. Models can either be conforming or non-conforming with a bound surface mesh, although when using non-conforming FE models with cuboid elements, particularly for GPU simulation, various optimisations can be made for increased computational performance, and such models are also easier to create and offer increased stability during simulations.

We have proposed a model creation process for automatically creating animatable non-conforming hexahedral FE simulation models, to which the object surface meshes are bound, has been presented. This process involves using samples within voxels to discretising the volume enclosed by surface meshes, possibly with internal surfaces, and calculate the element material properties and
muscle properties for each overlapping muscle. Muscle fibre directions are calculated using fibre field functions created from NURBS volumes, and boundary conditions are also computed by computing rigid and sliding nodes to constrain models.

Further details regarding the CUDA TLED FE system introduced in the 12 month report have been described [War11], which now contains functionality to simulate 8-node under-integrated hexahedral elements with hourglass control. An enhanced muscle contraction model has also been implemented to generate both active and transversely-isotropic passive muscle stresses, along with a component to simulate sliding nodes that are bound by and slide over a surface mesh.

Using soft-tissue block and facial models created using our model creation process with upscaled material densities for increased stability, simulations have been performed, which can simulate wrinkling on a soft-tissue block model (when the simulation model conforms to the outer surface mesh), and gross forehead movement due to frontalis muscle contraction. However, problems currently occur when using non-conforming models containing sharp steps (e.g. with curved models or models that aren’t aligned with the global axes when voxelising). Current work is focussing on rectifying these, for example, by smoothing the outer elements of simulation models. Further work will then focus on simulating additional material behaviour, experimenting with different soft-tissue block and facial models, and evaluating the system.

Various further improvements could be made, such as by enabling creation of multi-resolution simulation models with small elements around the outer surfaces, and larger elements in the middle of the model. A mixture of cuboid and tetrahedral elements could also be used to produce more accurate models that conform relatively well with the surface meshes, behave realistically, and can be simulated efficiently. The main focus at the moment, however, is the simulation of forehead wrinkles using models created with the current model creation process.

A plan for the remainder of the project has been devised. The main tasks to be completed are listed below, and the plan is illustrated in Figure 9.1.

1. Further Implementation and Experimentation

   (a) Complete implementing the mesh smoothing component. Use this to produce smoothed soft-tissue block and facial models, and perform simulations to test whether wrinkles are produced without artefacts.

   (b) Implement any further functionality necessary for wrinkle simulation. This may include, for example, functionality to better capture the properties of different skin layers when creating non-conforming models, or implementing shell element functionality (discussed in Section 7.1). Further research alongside this stage may also reveal other potential solutions.

   (c) Experiment using different models, such as soft-tissue block models at different angles to the global axes, and with different curvature, for wrinkle simulation to determine whether, and in which cases, wrinkles can be successfully simulated.

   (d) Implement additional material models and behaviour, such as the Ogden material model and transversely isotropic behaviour using a preferred fibre direction (discussed in Section 7.1).

   (e) Experiment using more complex models, for example, with multiple opposing muscles.

   (f) Complete the simple model of the forehead, which should include at least the frontalis, procerus and corrugator supercilii muscles, and these should be tweaked until realistic-looking animations can be produced.

   (g) Experiment with forehead gross movement and wrinkle simulation.

   (h) Experiment with material properties to simulate dynamic forehead wrinkles for different aged soft tissue. This is probably the only experimentation that will be done with different aged soft-tissue simulations. Other functionality, such as integrating plasticity effects into the material models to better simulate soft-tissue aging (as previously proposed) is now probably beyond the scope of this project.

2. Evaluation

   (a) Collect data, produce animations and design questionnaire for evaluation.

   (b) Perform quantitative evaluation by comparing the captured data and measurements to the simulations.
3. Write Up Thesis

It is expected that all practical work, including evaluations, should be finished around the 30 month mark. If further complications arise, for example, using non-conforming models to simulate wrinkles, it is possible that models which conform to the outer surfaces of the forehead region could be used (e.g. by deforming a simple flat model of the forehead, like with the curved soft-tissue block model in Section 6). With this approach, the models are essentially more complex conforming soft-tissue block models, which, as shown by Figure 6.1 in Section 6, can be used to simulate wrinkling behaviour. Such models could also possibly be attached to the existing facial model.

Some further tasks will also be completed alongside the main project:

- Writing papers or preparing presentations for, and attending conferences, workshops and seminars. Potential research papers or posters that could be written and submitted in the near future include:
  - A survey paper for physically-based facial animation, based on an updated version of the literature review presented in the 12 month report [War11].
  - A description the updated version of the model creation process to automatically produce simulation-ready non-conforming hexahedral FE models.
  - When the skin wrinkling issues have been sorted, a description the CUDA TLED FE simulation system, which can be used with our model creation process for efficient simulation of skin wrinkling. This will include examples of soft-tissue block and forehead wrinkles.
  - A comparison of wrinkling on different foreheads (e.g. foreheads of different sizes and curvatures) using real data and simulations using our simulation system.

- Submitting applications for any fellowships. The NVIDIA Graduate Fellowship will be applied for again this year, for which the deadline is 15th January 2013.

- DDP-related tasks, such as demonstrating, which includes two modules totalling 54 hours in the upcoming semester.
Conclusion and Project Plan

*** Further Implementation and Experimentation ***

- Complete the mesh smoothing component
- Implement further necessary functionality for wrinkle simulation
- Experiment using models of different curvature for wrinkle simulation
- Implement additional material models and behaviour
- Experiment using models with multiple opposing muscles
- Complete the simple model of the forehead
- Experiment with forehead gross movement and wrinkle simulation
- Experiment with material properties for different aged soft tissue

*** Evaluation ***

- Collect data and produce animations for evaluation
- Perform quantitative evaluation
- Perform qualitative evaluation
- Write Up Thesis

Figure 9.1: A plan of the remaining project tasks to be completed.
References


A DDP Progress

Further to the DDP progress reported in the 12 month report [War11], various activities have been completed:

- Attended and presented a paper at WSCG 2012\(^3\) [WM12].
- Presented research to the graphics research group\(^4\).
- Submitted an application for the NVIDIA Graduate Fellowship Program 2011/12.
- Completed the following modules as part of the DDP:
  - CIC6003 (Techniques for High Performance Computing) to learn more about parallelisation, and using parallel programming frameworks, such as MPI and OpenMP, on supercomputers.
  - FCE6100 (Research Ethics and Integrity) to learn about the importance of ethics and discuss different ethical issues that may arise in computer science and engineering research.
- Several activities relating to teaching and demonstrating have been completed:
  - Completed laboratory demonstrating activities and assignment marking for the following modules:
    * 40 hours for module COM3503/4503/6503 (3D Computer Graphics)
    * 50 hours for module COM1004 (Web and Internet Technology)
  - Currently attending STA workshops on laboratory demonstrating and assessment and feedback to help improve demonstrating-related skills, and assignment marking. Attendance at these is required to continue demonstrating.
- Continuously maintaining a personal website detailing any research, experiments and demos that have been made.
- Made contact with experts in different fields:
  - Contacted Dr Zeike Taylor from the Department of Mechanical Engineering at the University of Sheffield to discuss issues relating to the stability of explicit FE simulations.
  - Met Dr Adrian Jowett from the School of Clinical Dentistry at the University of Sheffield to discuss the anatomy of soft tissue in the forehead region, such as the attachments of the frontalis muscle.
- The following demonstrating hours for both lab classes and assignment marking have been allocated for the current academic semester:
  - 24 hours for module COM3503/COM4503/COM503 (3D Computer Graphics)
  - 30 hours for module COM1004 (Web and Internet Technology)
