Unified Skinning of Rigid and Deformable Models for Anatomical Simulations

Ian Stavness*[1], C. Antonio Sánchez[2], John Lloyd[2], Andrew Ho[2], Johnty Wang[3], Sidney Fels[2], Danny Huang[1]
[1]University of Saskatchewan [2]University of British Columbia [3]McGill University

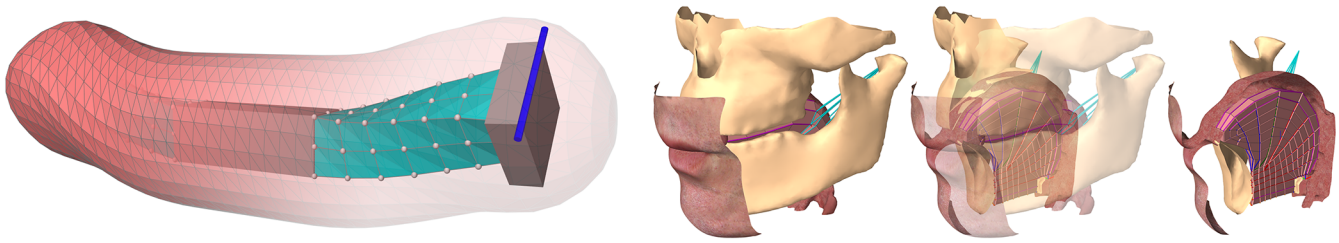

Figure 1: *Our unified geometric skinning method for rigid and deformable bodies with two-way force coupling (left), is well suited for anatomical models that have a mix of hard and soft tissues, such as dynamic simulations of the tongue, jaw, skull, and vocal tract (right).*

Abstract

We propose a novel geometric skinning approach that unifies geometric blending for rigid-body models with embedded surfaces for finite-element models. The resulting skinning method provides flexibility for modelers and animators to select the desired dynamic degrees-of-freedom through a combination of coupled rigid and deformable structures connected to a single skin mesh that is influenced by all dynamic components. The approach is particularly useful for anatomical models that include a mix of hard structures (bones) and soft tissues (muscles, tendons). We demonstrate our skinning method for an upper airway model and create first-of-its-kind simulations of swallowing and speech acoustics that are generated by muscle-driven biomechanical models of the oral anatomy.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation;

Keywords: skinning, physics-based simulation, finite-element modeling, airway, swallowing, speech production

1 Introduction

Skinning is a widely-used computer animation technique for separating the geometric fidelity of a model's appearance from the kinematic fidelity of its dynamics. Skinning allows a character rig with very few degrees of freedom (such as a skeleton with 14 joints) to drive the deformations of a very high fidelity geometric model (such as a body surface mesh with millions of polygons).

Limitations of geometric skinning, such as unrealistic bulging, pinching and twisting, especially near the location of joints, such as the elbow, have led researchers to pursue physically-based skinning techniques. These add dynamics to the skin mesh in order to improve skin deformations, handle collisions, and generate secondary motions such as bulging and wrinkling [McAdams et al. 2011]. Physically-based skinning, however, typically entails a large increase in a model's degrees-of-freedom, making it computationally expensive. As a compromise between physical realism and computational efficiency, researchers have proposed a reduced technique: embed a high resolution geometric mesh within a coarse finite-element model [Capell et al. 2002].

We extend this intermediate-fidelity skinning approach to work seamlessly for both rigid bodies and finite-element models that drive a geometric skinning mesh. This gives significant flexibility to modelers and animators to choose the appropriate level of fidelity for the dynamics underlying their models, while maintaining a detailed surface geometry. Our approach can capture large movements with a rigid rig representing the skeleton, include coarse and fast deformable structures for selective soft-tissue structures, and have a high-fidelity geometric skin model influenced by both rigid and deformable structures simultaneously (Figure 1).

The upper airway presents unique modeling challenges and is particularly well-suited to our approach of intermediate fidelity skinning. The mucosa lining of the mouth and throat forms an internal skin surface that covers a number of hard and soft anatomical structures, including the gums, hard palate, tongue, soft palate, and pharynx. This internal skin is similar to the external skin of the body, but it is enclosed by surrounding structures, as opposed to surrounding underlying structures.

Applications in upper airway modeling, such as simulations of swallowing, breathing, and speaking, all require a closed water-tight geometry that represents the mucosa lining of the mouth and throat. The mucosa lining forms a boundary condition for the flow of air, fluid and food. By representing this as a skinned mesh, we can perform simulations in the domain of the air-space, such as the fluid dynamics of a liquid food bolus during swallowing, or airflow during speaking. This alleviates the need to have water-tight connections between the underlying dynamic structures. The complete airway can then be driven by a select few dynamic components that have the greatest affect on airway shape, affecting flow for both speech and swallowing processes.

The contributions of our work include:
1) A unified geometric skinning technique over both rigid body and deformable body structures;
2) A general dynamic attachment paradigm that allows forces (including contact) to be transmitted back and forth between to the skin mesh and underlying dynamic components;
3) A demonstration of airway skinning for simulating swallowing of a fluid food bolus with a muscle-driven tongue model; and
4) A demonstration of airway skinning for simulating speech vowel production acoustics in a dynamic vocal tract.

*e-mail:stavness@gmail.com

SIGGRAPH Asia 2014, December 03 – 06, 2014, Shenzhen, China.

ACM 978-1-4503-2895-1/14/12
http://dx.doi.org/10.1145/2669024.2669031

2 Related work

Geometric skinning most commonly uses an explicit surface mesh and blends vertices between the rigid bodies of a skeleton. Linear blending [Thalmann et al. 2004] is efficient, but has well known visual artifacts around bending joints (pinching, bulging, "candy-wrapper" twisting) that can be partly alleviated by more sophisticated blending [Kavan et al. 2008]. With carefully selected blending weights, skinning approaches can achieve compelling visual results, however geometric skinning of a rigid skeleton still has limitations in generating physical realistic responses during animations.

To capture richer deformations of skin and other surfaces, researchers have proposed physically-based skinning methods. These techniques reduce geometric artifacts of skinning and are also used to capture secondary motion artifacts (bulging, wrinkling, etc.) that add to the physical realism of a character's animation. Most commonly, finite-element methods are employed to add elasticity to the skin [McAdams et al. 2011] and sliding of skin over subcutaneous structures [Li et al. 2013]. The additional realism afforded by physically-based skinning comes at the cost of greater computational complexity as the dynamics involve many more degrees-of-freedom. To trade off between physical realism and computational cost, an intermediary approach is to embed a high resolution surface mesh into a lower resolution physics-based model.

Mesh embedding reduces the degrees of freedom in the physical model while maintaining a high fidelity surface mesh [Capell et al. 2002], making it a popular choice. The general approach is to use a set of control points to govern the motion of an encapsulated volume. The control lattice is typically provided by a finite-element model, allowing more realistic deformations determined by a set of underlying material properties [Capell et al. 2002; Teschner et al. 2004]. The internal deformation field is then provided by the finite-element shape-functions. Other low-resolution deformable techniques have also been employed, such as the frame-based meshless technique of [Faure et al. 2011]. With this method, the control points are rigid-body frames. Skinning weights are proximity-based, and the geometry is interpolated using frame-based skinning techniques such as dual-quaternion blending [Kavan et al. 2008]. It is often necessary to allow interactions (such as collisions) with the embedded geometries rather than with the coarser grid [Nesme et al. 2009]. Thus, a method of distributing forces back to the control lattice is required. This has been reported independently for point-like nodes of a finite-element model [Kaufmann et al. 2009] and for rigid-body frames [Faure et al. 2011]. In this work, we provide a unified framework for distributing forces back to all control components, including both point-line nodes and frames.

3 Methods

3.1 Simulation and attachments in ArtiSynth

The simulations in this paper are performed using ArtiSynth (www.artisynth.org), a modeling platform that supports combined multi-body and finite element simulation, together with contact and constraints. Multi-body finite-element coupling is achieved by (a) implicit-integration, and (b) an attachment mechanism that allows dynamic components (e.g., particles or rigid bodies) to be connected to each other. Full details are given in [Lloyd et al. 2012].

Components can be physically attached together by making the position $\mathbf{q}_a$ of the attached component a function of one or more *master* components, whose collective positions are given by $\mathbf{q}_m$:

$$\mathbf{q}_a = f\left(\mathbf{q}_m\right). \tag{1}$$

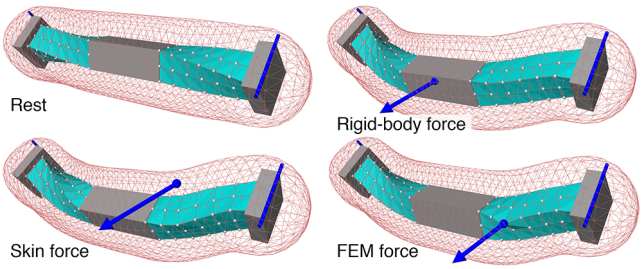

Figure 2: *External force (blue arrow) applied to a rigid body (grey), finite-element model (cyan) or skin mesh (pink wireframe).*

By differentiating with respect to time and linearizing, we can obtain an approximate relationship between the velocities $\mathbf{u}_a$ and $\mathbf{u}_m$ of the attached component andits masters:

$$\mathbf{u}_a = \mathbf{G}\,\mathbf{u}_m, \quad \mathbf{G} \equiv \partial f / \partial \mathbf{q}_m \tag{2}$$

ArtiSynth employs (2) in its per-step solution for all dynamic component velocities, and it then applies (1) to correct for numeric drift.

3.2 Unified skinning for rigid body and finite-element components

Our skinning framework is based on the previous attachment mechanism, in which each skinned point is treated as a *virtual* dynamic component that is attached to one or more master components, which can be either 3-DOF points, such as finite-element nodes, or 6-DOF frames, such as rigid body coordinates. The position of each vertex, $\mathbf{q}_v$, is given as a weighted sum of contributions from each master component:

$$\mathbf{q}_v = \mathbf{q}_{v0} + \sum_{i=1}^{M} w_i f_i(\mathbf{q}_m, \mathbf{q}_{m0}, \mathbf{q}_{v0}) \tag{3}$$

where $\mathbf{q}_{v0}$ is the initial position of the skinned point, $\mathbf{q}_{m0}$ is the collective rest state of the masters, w_i is the skinning weight associated with the ith master component, and f_i is the corresponding blending function.

For a point master (e.g. FE node), the blending function is the displacement of the single point:

$$f_i(\mathbf{q}_m, \mathbf{q}_{m0}, \mathbf{q}_{v0}) = \mathbf{q}_i - \mathbf{q}_{i0}. \tag{4}$$

For frames, we currently allow for linear or dual-quaternion linear blending. Let the position/orientation of the control frame be encoded as a dual-quaternion $\mathbf{q}_i$. For linear blending,

$$f_i(\mathbf{q}_m, \mathbf{q}_{m0}, \mathbf{q}_{v0}) = \hat{\mathbf{q}}_i \mathbf{q}_{v0} \hat{\mathbf{q}}_i^{-1} - \mathbf{q}_{v0}, \quad \hat{\mathbf{q}}_i = \mathbf{q}_i \mathbf{q}_{i0}^{-1}. \tag{5}$$

The quaternion product $\hat{\mathbf{q}} = \left(\mathbf{q}\mathbf{q}_0^{-1}\right)$ is the relative control frame transform, which is applied to initial vertex location $\mathbf{q}_{v0}$. For dual-quaternion skinning, the expression is more complex:

$$f_i(\mathbf{q}_m, \mathbf{q}_{m0}, \mathbf{q}_{v0}) = \tilde{\mathbf{q}}_i \mathbf{q}_{v0} \tilde{\mathbf{q}}_i^{-1} - \mathbf{q}_{v0} \tag{6}$$

$$\tilde{\mathbf{q}}_i = \frac{\mathbf{q}_i \mathbf{q}_{i0}^{-1}}{\left\| \sum_{j=1}^{M} w_j \mathbf{q}_j \mathbf{q}_{j0}^{-1} \right\|}. \tag{7}$$

In this system, the contribution of the ith control frame is scaled such that the net transform has a unit norm. For complete details on dual-quaternion skinning, refer to [Kavan et al. 2008].

The position equation (3) can be differentiated to yield the linearized velocity relationship. For point masters, the positional relationships (4) are linear, so this is trivial. For frame masters, it is necessary to linearize (5) or (7) to compute $\mathbf{G}$.

3.3 Force distribution and collisions

To provide two-way coupling between the skinned mesh and its underlying dynamic components, it is necessary to propagate forces acting on the skinned vertices back to the dynamic masters. Given the relationship (2), the principle of virtual work dictates that generalized forces $\mathbf{f}_a$ acting on the attachment propagate back to the master forces $\mathbf{f}_m$ according to

$$\mathbf{f}_m = \mathbf{G}^T \mathbf{f}_a \tag{8}$$

This means that the velocity relationships presented in Section (3.2) can be used, in transposed form, to send forces from the points back to the underlying master points or frames. An example of this is shown in Figure 2, where a point-force is applied to a single location on the skinned mesh.

The velocity relation can also be used to handle collisions. Within ArtiSynth, collisions between deformable meshes are handled using unilateral constraints of the form

$$\mathbf{N}\mathbf{u}_c \geq 0, \tag{9}$$

where $\mathbf{u}_c$ is the assembled velocities of the mesh vertices associated with the contact (typically a vertex and a face). For a skinned mesh, (2) gives each vertex velocity in terms of the velocities of its underlying master components, and so substituting this into (9) yields a unilateral constraint acting on the master components:

$$\mathbf{N}'\mathbf{u}_m \geq 0, \tag{10}$$

These constraints are incorporated directly into the system, and solved simultaneously with the rest of the dynamics.

The complexity of the skinning algorithm is $\mathcal{O}(n)$ where n is the number of vertices in the skin mesh. In practice, the time spent on skinning is negligible compared to that of the dynamics simulation.

4 Applications to upper airway modeling

We have applied our skinning framework to a 3D biomechanical model of the upper airway. The model is composed of the main anatomical structures that affect the size and shape of the airway. The overall model is pictured in Figure 1 and its constituents are isolated in Figure 3. The model includes finite-element meshes for the face, tongue, and soft-palate, as well as rigid meshes for the jaw, skull, larynx, and pharynx. These model components are dynamically coupled and movements are generated by activating muscles in the model. Individual components can also be kinematically-driven where appropriate to reduce the dynamic degrees-of-freedom in the model and increase simulation speed.

The airway mesh that represents the mucosa lining of the mouth and throat was segmented from CT data for the subject upon which the model is based. It is skinned to the anatomical structures and blending weights are chosen by the distance between surfaces.

4.1 Swallowing simulation

The oral phase of swallowing involves a coordinated movement of the tongue and soft-palate. We simulated these movements with a stationary jaw and hard-palate, a kinematically-driven soft-palate and a muscle-driven FE tongue. Muscles fibers embedded within the tongue were activated in order to track the desired movement of three points on the upper surface of the tongue (with ArtiSynth's inverse simulation tool [Lloyd et al. 2012]).The soft-palate was set to move from a low to high position to close off the nasal passage as the fluid bolus moves to the back of the throat.

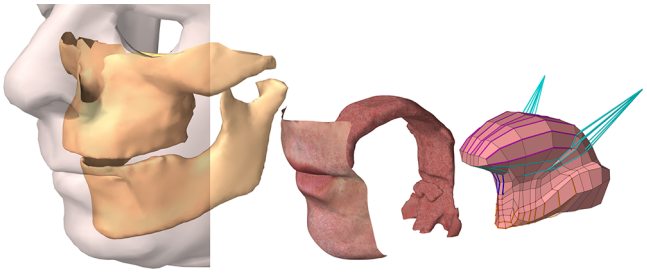

Figure 3: *Upper airway model components: face and jaw models, airway skin mesh, finite-element tongue model.*

SPH simulation We simulated a liquid bolus in our swallowing animations with the SPH formulation described in [Ho et al. 2014]. SPH is a Lagrangian, mesh-free method where the fluid domain is represented by discrete particles. Here the incompressible, viscous, iso-thermal Navier-Stokes equations are solved. SPH simulation requires a water-tight boundary to contain the fluid bolus. Without a proper boundary, leakage of the bolus may occur, for example between the lateral tongue and the upper teeth. Our skinning formulation provides a airway boundary with high spatial resolution and little computational effort. In our swallowing simulation, a 5mL fluid bolus is simulated with 5213 particles.

Results The skinned airway successfully contained the bolus in the oral cavity at the beginning of the sequence (Figure 4, left column) and permitted the bolus to flow into the pharynx (Figure 4, middle and right columns). The airway skin provided a pharyngeal wall without the need for an explicit FEM of the pharynx. The movement of the bolus for the oral phase looks qualitatively similar to that of normal swallows (Figure 4, lower panel).

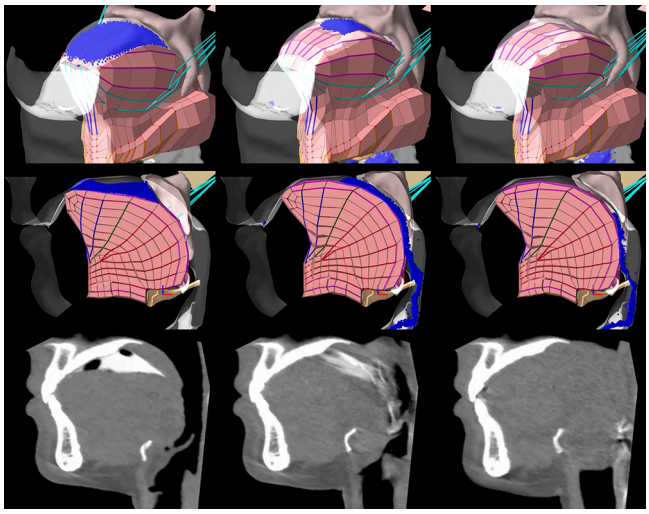

Figure 4: *Swallowing simulation results: 3D (top row) and lateral (middle row) views of airway skin mesh deformed by tongue and soft-palate with SPH fluid bolus (blue). CT images of a real swallow (bottom row, courtesy of Eiichi Saitoh, Fujita Health Univ.).*

4.2 Speech acoustics simulation

Speech sounds are generated by vibrations of the vocal folds inside of the larynx that resonate through the vocal tract and emanate through the mouth and nose. Changes in sound characteristics occur both due to changes in the vocal folds (tightening, opening, et al.) and changes to the shape of the vocal tract under muscle con-

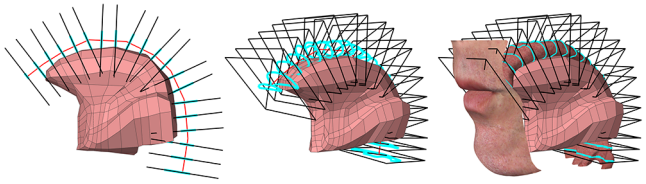

Figure 5: *Cross-sectional cut planes (black) intersect with airway skin mesh (cyan contours). The area enclosed by each contour parameterize a source-filter based acoustics simulation.*

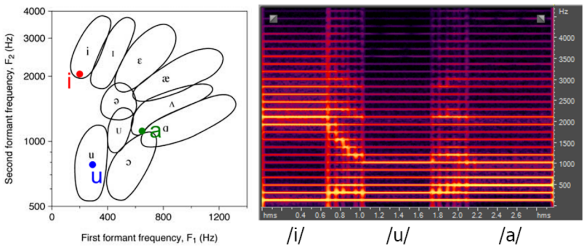

Figure 6: *Airway acoustics simulation results: the formant freq. of synthesized vowels relative to typical human phonation (left). Spectrogram of resulting transitions between the vowels (right).*

trol. To synthesize acoustics with our dynamic airway model, we treat the sound source (glottis) separately from the resonating tube (vocal tract), a common approach dating to the 50s [Fant 1970].The resonating tube is represented as a transfer function defined by the cross-sectional areas of 20 segments of the vocal tract. The glottal sound source is a two-mass model of the oscillations of the vocal folds. We couple this source to the filter model derived from the cross-sectional areas and solve a 1D implementation of the Navier-Stokes equations [van den Doel and Ascher 2008].

As can be seen in Figure 5, from our upper airway model, we can calculate, in real time, cross-sectional areas at regular intervals along the vocal tract from the skinning mesh attached to the various muscle-activated articulators. These cross-sectional areas are then sent to the source-filter synthesizer so that as the upper airway articulators deform the airway mesh the sound is modulated.

Results Using this approach, we simulated vowel postures by moving the tongue to expected locations of the tongue for the cardinal /i/ , /u/, and /a/ vowels. The first and second formant frequencies were within human range for /u/ and on the border of the range for /i/ and /a/ (Figure 6, left). We also used moving data from the dynamic simulation between the static postures, and produced an /i-u-a/ transition trajectory that is shown in right panel of Figure 6. The results compare favorably with established formant frequencies [Fant 1970], though, considerable effort will be needed to tune and investigate the perceptual performance of our approach.

5 Limitations and future directions

Our skinning framework has a few limitations that we plan to address in future work. We currently do not handle self collisions in the airway mesh, however, if self collisions were found with a standard detection method they could be resolved with our existing collision handling approach. We observed large amounts of stretching of the airway mesh when the tongue moved backward and underwent large deformations, such as in backward motion of the tongue tip, which generates significant stretching of the airway mesh. A sliding approach may be more appropriate for tongue-airway skinning, as proposed in [Li et al. 2013].

Our airway simulations are intended as a proof-of-concept of integrating muscle-driven biomechanical models that surround the upper airway with fluid and acoustic phenomena that occur within it. We have only considered the oral phase of swallowing, whereas a full swallow, including bolus transport to the esophagus, will require active models of the laryngeal structures that close off and protect the airway and constriction of the pharyngeal wall. Our speech simulation has many directions for improvement in order to generate more appealing acoustics, including coupling the glottal source to the filter and an aeroacoustic model of the vocal folds.

6 Conclusions

In this paper we proposed a novel geometric skinning approach that unifies geometric blending for multibody models with embedded surfaces for finite-element models. The resulting skinning approaches provides flexibility for modelers and animators to select the desired dynamic degrees-of-freedom through a combination of coupled rigid and deformable bodies and specify a single skin mesh influenced by all dynamic components. This approach is particularly well suited for anatomical simulations because biological structures include a mix of both hard and soft tissues. We demonstrated the utility of unified skinning for simulating a closed-water tight airway mesh that is deformed by the anatomical structures of the vocal tract. This airway skin provides a convenient and efficient boundary for performing fluid and acoustics simulations in the airway. Our simulation results illustrate the first-of-its-kind swallowing and speech acoustics simulations that are generated by muscle-driven biomechanical models of the oral anatomy.

References

CAPELL, S., GREEN, S., ET AL. 2002. Interactive skeleton-driven dynamic deformations. In *ACM SIGGRAPH*, 586–593.

FANT, G. 1970. *Acoustic Theory of Speech Production*.

FAURE, F., GILLES, B., ET AL. 2011. Sparse meshless models of complex deformable solids. In *ACM SIGGRAPH*.

HO, A. K., ET AL. 2014. A 3d swallowing simulation using smoothed particle hydrodynamics. *CMBBE: Imag & Vis in press*.

KAUFMANN, P., MARTIN, S., ET AL. 2009. Flexible simulation of deformable models using ... FEM. *Graph Models 71*, 153–167.

KAVAN, L., COLLINS, S., ET AL. 2008. Geometric skinning with approximate dual quaternion blending. In *ACM SIGGRAPH*.

LI, D., SUEDA, S., NEOG, D. R., AND PAI, D. K. 2013. Thin skin elastodynamics. In *ACM SIGGRAPH*.

LLOYD, J. E., ET AL. 2012. Artisynth ... modeling toolkit combining multibody and FE simulation. Springer, 355–394.

MCADAMS, A., ET AL. 2011. Efficient elasticity for character skinning with contact and collisions. In *ACM SIGGRAPH*.

NESME, M., ET AL. 2009. Preserving topology and elasticity for embedded deformable models. In *ACM SIGGRAPH*.

TESCHNER, M., HEIDELBERGER, B., ET AL. 2004. A versatile and robust model for geometrically complex deformable solids.

THALMANN, N. M., CORDIER, F., ET AL. 2004. Modeling of bodies and clothes for virtual environments. In *Cyberworlds*.

VAN DEN DOEL, K., AND ASCHER, U. 2008. Real-time numerical solution of Webster's equation ... *IEEE Trans Audio Speech 16*.